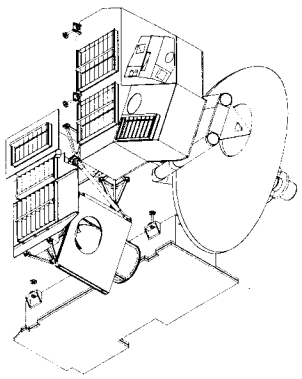
JPL D-16159
EOS MLS DRL 601 (part 3)
ATBD-MLS-03

## Earth Observing System (EOS)

## Microwave Limb Sounder (MLS)

# EOS MLS Retrieval Processes Algorithm Theoretical Basis



$$\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\mathrm{y}}^{-1}\mathbf{K} = \begin{bmatrix} \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & 0 & 0 & 0 \\ \times & \times & \times & \times & \times & 0 & 0 \\ \times & \times & \times & \times & \times & \times & 0 \\ \times & 0 & \times & \times & \times & \times & \times \\ \times & 0 & 0 & \times & \times & \times & \times \\ \times & 0 & 0 & 0 & \times & \times & \times \end{bmatrix}.$$

Nathaniel J. Livesey,
Dong L. Wu.

**Version 1.1**
October 15, 1999

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California, 91109-8099

# Release Record

| Version | Date | Comments |
|---|---|---|
| 1.0 | 15 January 1999 | Initial version. |
| 1.1 | 15 October 1999 | Released following formal review of Version 1.0 by NASA board reviewing the EOS CHEM Algorithm Theoretical Basis Documents. This document received top grade of 'A' from the review board. The board recommendations which apply to this document are italicized below, followed by the (non-italicized) responses of the MLS team.<br>• *All aspects of the forward model should be examined ...* As outlined in the overview document, a separate forward model ATBD is in preparation.<br>• *A more complete and realistic evaluation of the retrieval algorithm CPU time requirements should be performed.* The concern during the review was the requirements for forward model CPU time, this will be discussed in the forward model ATBD. The CPU requirements for the inverse model are discussed in detail in Section 7.4 of this document. More detailed estimates will be given in later versions of this document, based on early versions of the production code.<br>No changes to the document were needed as a result of these recommendations.<br>Changes from Version 1.0 are described below, reflect expected progress and correction of a few minor errors.<br>• The 'secondary' geophysical products of Version 1.0 have now been defined as 'standard' data products. This resulted in updates to Table 1.1.<br>• Also in Table 1.1 the product designation has been changed to use dashes rather than underscores, to reflect the implementation in software.<br>• Added definition of $g_0$ in Section 3.4.<br>• Modified Figure 3.6 to include effects of refraction. Caption modified accordingly.<br>• Chapter 6 largely rewritten to reflect new methodology.<br>• Modified Section 7.5 as it is likely that the forward model radiance data will be produced and archived routinely.<br>• Corrected Section C.2 to use $\log_2$ rather than ln.<br>• Modified Table B.1 to use dashes rather than underscores, as in Table 1.1.<br>• Minor grammatical corrections. |

# Contents

# Chapter 1
# Introduction

## 1.1 The Microwave Limb Sounder experiment

EOS MLS is a successor to the MLS experiment (Barath et al. 1993) that formed part of the Upper Atmosphere Research Satellite (UARS), launched in September 1991 (Reber 1993; Reber et al. 1993). The instrument is designed to study aspects of the chemistry and dynamics of the atmosphere, from the upper troposphere to the lower mesosphere. The microwave heterodyne technique is employed to observe thermal microwave emission from the Earth's limb in several spectral bands, designed to characterize emission from $O_2$ (used to obtain temperature and pressure information), $O_3$, $H_2O$, ClO, HCl, $HNO_3$, $N_2O$, CO, OH, $SO_2$, BrO, HOCl, $HO_2$, and HCN.

EOS MLS is one of four instruments on the EOS CHEM spacecraft. The launch of the CHEM platform is planned for late 2002. CHEM will fly in a 98° inclined orbit, at a nominal height of 705 km, with a nominal orbital period of 98.9 minutes.

## 1.2 The aims of this document

This document describes the theoretical basis for the 'retrieval' algorithms to be used in the routine processing of data from the MLS instrument. The task of the retrieval algorithms is to convert calibrated measurements of thermal limb emission from MLS into estimates of geophysical parameters such as temperature and composition.

The MLS calibrated radiance observations are known collectively as Level 1B data. Level 2 data describe retrieved geophysical parameters at the nominal footprint of the MLS instrument. These Level 2 data are produced using the algorithms described in this document. Level 3 data describe geophysical parameters on regular latitude/longitude grids. Most of the Level 3 products are generated by the application of gridding algorithms to the Level 2 data.

Some of the species MLS is designed to measure have very low mixing ratios and/or weak emission lines. This leads to a comparatively poor signal to noise ratio for the corresponding MLS radiance observations. The best estimates of the abundance of such species are obtained by using retrieval algorithms to derive quantities such as daily zonal means, monthly maps, or similar averaged quantities. Such datasets are by definition Level 3 data. Because the algorithms used to produce these data necessarily use retrieval theory, they are also described in this document.

## 1.3 Related Algorithm Theoretical Basis Documents

An overview of the MLS instrument and data processing operations is given in Waters 1999. The algorithms used in the Level 1 processing to calibrate the raw observations of microwave radiance made by the instrument are described in Jarnot 1999. An important component of the Level 2 processing algorithms is the *forward model*, the theoretical basis for this aspect of the data processing is given

**Table 1.1:** A list of the geophysical products produced by the algorithms described in the document. For a definition of the terms used see Waters 1999. The products listed in bold type are the standard MLS data products, those in non-bold type are additional 'diagnostic' products produced by the algorithms. All the products, both standard and diagnostic, are produced using the same algorithms. The 'diagnostic' products are produced for special scientific and/or diagnostic purposes. Products marked † are best processed using the 'noisy' products algorithm to produce Level 3 data for some or all of their vertical range, although Level 2 data for these may be produced routinely. The quantities marked ‡ are produced using algorithms that are not based on retrieval theory, but on analysis of other retrieved products.

**Geophysical products:**

| | | |
|---|---|---|
| **TEMPERATURE** | **HNO3** | BrO-625† |
| TEMPERATURE-118-a | **O3** | BrO-650† |
| TEMPERATURE-118-b | O3-190 | **CO** |
| TEMPERATURE-240 | O3-240-4 | **HOCl**† |
| **GEOPOTENTIAL-HT** | O3-240-9 | **HCN**† |
| GEOPOTENTIAL-HT-118-a | O3-640 | **SO2** |
| GEOPOTENTIAL-HT-118-b | O3-STRAT-COLUMN | SO2-204 |
| GEOPOTENTIAL-HT-240 | **HCl** | SO2-640 |
| **H2O** | **ClO** | **CIRRUS-ICE**‡ |
| H2O-TROP-118 | ClO-240 | CLOUD-EXT-COEFF-118 |
| H2O-TROP-190 | ClO-640 | CLOUD-EXT-COEFF-190 |
| H2O-TROP-240 | **N2O** | CLOUD-EXT-COEFF-240 |
| H2O-TROP-640 | N2O-190 | CLOUD-EXT-COEFF-640 |
| H2O-TROP-2.5T | N2O-640 | CLOUD-EXT-COEFF-2.5T |
| **REL-HUM-TROP** | **OH**† | TROPOPAUSE-PRESSURE‡ |
| REL-HUM-TROP-118 | OH-2514-H† | |
| REL-HUM-TROP-190 | OH-2514-V† | |
| REL-HUM-TROP-240 | OH-2510-H† | |
| REL-HUM-TROP-640 | OH-2510-V† | |
| REL-HUM-TROP-2.5T | **HO2**† | |
| | HO2-650† | |
| | HO2-660† | |
| | **BrO**† | |

in Read et al. 1999. The precisions to be expected from the EOS MLS observations of geophysical parameters are given in Filipiak 1999.

## 1.4 EOS MLS data products for which this document applies

The algorithms described here are used in the production of all the EOS MLS geophysical products. For a complete list of the products see Table 1.1. The majority of the products in the table will be produced using the retrieval algorithms described in Chapters 4 and 5. Cirrus ice products are described in Chapter 6. Tropopause pressure, and products such as column ozone are described in Appendix A.

# Chapter 2
# Overview of EOS MLS Level 2 data processing

Most of this document describes the algorithms used in the production of the Level 2 data. This chapter gives a brief overview of the components of the algorithms and the software which implements them.

## 2.1 The aims of retrieval theory

The methods used to convert remote measurements of radiation emitted by the atmosphere into estimates of geophysical parameters are known as *retrieval algorithms*. Retrieval theory is a well-established field in atmospheric science and is covered in the standard literature (Rodgers 1976; Rodgers 1990). A mathematical discussion of retrieval theory is given in Section 3.2. This section gives a brief outline of the fundamental principals involved.

In a retrieval algorithm a quantity known as the *state vector* is used to describe the current knowledge of the state of the atmosphere and relevant aspects of the instrumental calibration and state (known collectively as the measurement system). Typically the state vector is initialized with *a priori* information such as climatological datasets. Given the state vector, a *forward model* calculation can be applied to predict what radiances the instrument would observe, were the measurement system in the state described by the state vector. By comparing these predicted radiances with the radiances actually observed, and by making use of additional information provided by the forward model calculation (namely derivatives of radiance with respect to the state vector), the retrieval algorithm computes a better estimate of the state vector, i.e. one for which the predicted radiances will be closer to those observed. Often retrieval algorithms use an iterative approach to continuously refine the state vector until appropriate convergence has been achieved. For various reasons, including numerical stability, most retrieval algorithms include *virtual measurements*. These are typically *a priori* estimates of the contents of the state vector; usually, but not in all cases, the same estimates used as initial values in the iterations.

## 2.2 Structure of the Level 2 data processing software

The main components of the Level 2 data processing software are shown in Figure 2.1. The principal components are the retrieval and forward model calculations described above. Before these are invoked, operational meteorological data are combined with climatological datasets and *a priori* knowledge of the state of the instrument, in order to construct an initial value for the state vector. After the retrieval and forward model calculations are complete, an additional process produces other products (such as cloud parameters, tropopause pressure and stratospheric ozone column), and outputs the Level 2 data along with appropriate diagnostics.

## 2.3 Heritage of the MLS retrieval algorithms

The EOS MLS instrument is a successor to the MLS instrument that formed part of the Upper Atmosphere Research Satellite (UARS) mission, launched in September 1991. The data processing algo-

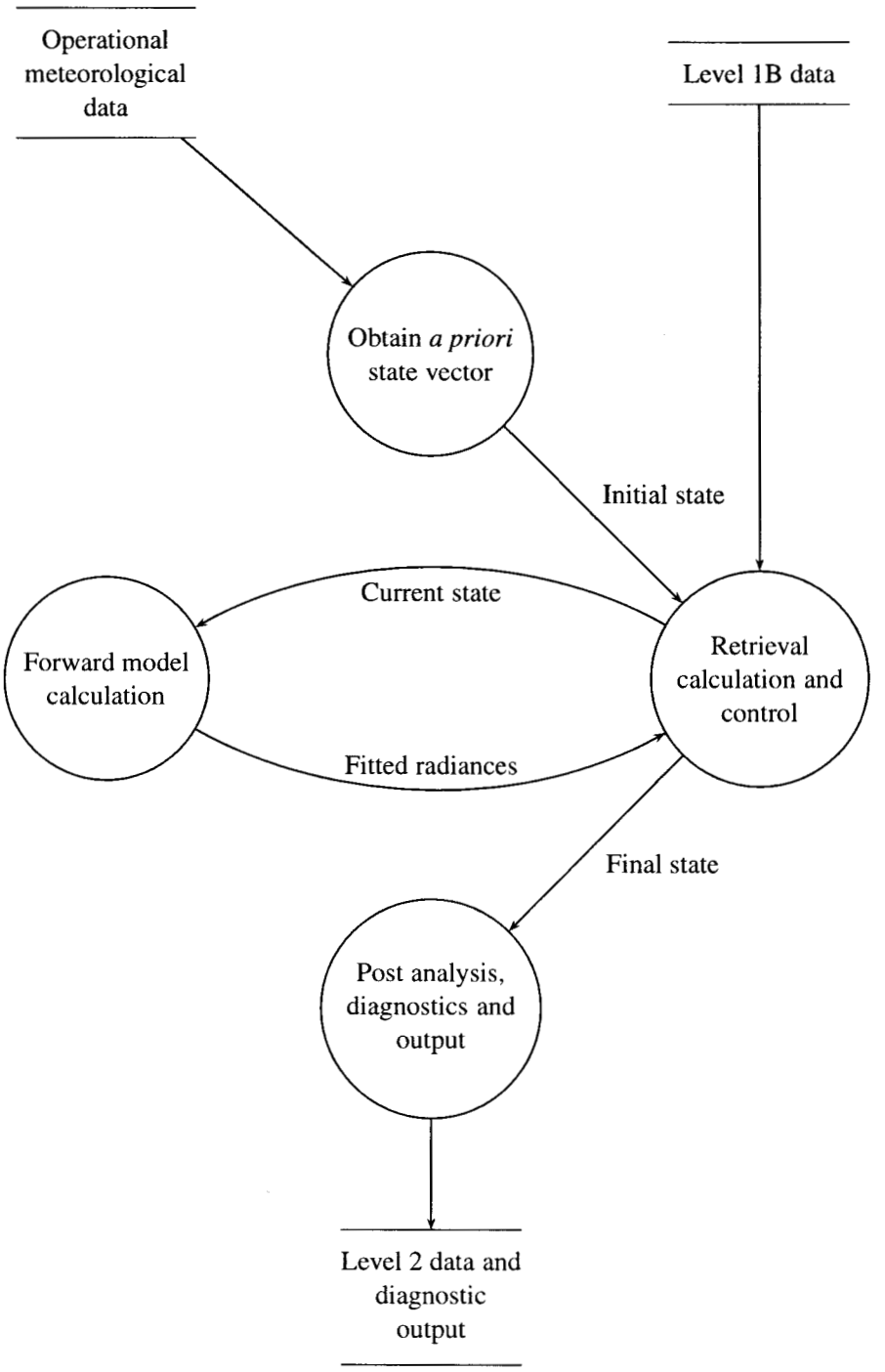



**Figure 2.1:** This figure shows the essential components and data flow for the EOS MLS Level 2 data processing software.

rithms envisaged for EOS MLS have essentially the same theoretical basis as those implemented for UARS MLS (notably those used to produce the latest version, v5, of the UARS MLS dataset). For information on the earlier versions of the MLS retrieval algorithms see Fishbein et al. 1996; Froidevaux et al. 1996 and Waters et al. 1996. The version 5 data processing algorithms for UARS MLS will be described in Livesey et al. (manuscript in preparation).

While the theoretical foundation of the EOS MLS retrieval algorithms is the same as that of the UARS algorithms, several aspects of the EOS MLS instrument design and intended scientific use differ from those of the UARS instrument, necessitating a somewhat different implementation of the algorithms. The major differences, in regards to the implication for algorithms are:

- The UARS instrument performs a 'stop and stare' scan, whereby the instrument views essentially the same region of the atmospheric limb for a brief period before moving onto a new region. The EOS MLS instrument performs a continuous scan, where the height of the limb path is continuously varied. This factor must be taken into account in the forward model calculation.

- The UARS MLS instrument observes limb emission in a direction perpendicular to the spacecraft flight direction, while the EOS MLS instrument observes emission from the region of the atmosphere directly ahead of the satellite. This geometry can, if properly exploited, yield significantly more information about the horizontal atmospheric variability along the measurement track.

- UARS MLS was designed to study processes mainly in the upper stratosphere. In recent years, scientific studies have become increasingly focused on the lower stratosphere and upper troposphere. The EOS MLS instrument targets this region specifically by the use of high bandwidth radiometers. Such observations however, can be affected by the presence of cloud (though not to the same degree as observations made using infrared/visible techniques). In addition, in these regions, the retrieval problem is typically more non-linear than in the upper stratosphere, due to the increased optical depth of the atmosphere at the wavelengths under consideration.

# Chapter 3
# The EOS MLS measurement system

This chapter outlines the physics behind the limb sounding technique and derives the algebraic expressions used in retrieval calculations. The aspects of the MLS instrument design that are relevant to the retrieval algorithms are summarized. The chapter then considers in detail how the state vector used in the MLS retrieval algorithms is constructed, and considers some implications of the MLS observation geometry.

## 3.1 The physics of limb sounding

Limb sounding of the atmosphere is a well established technique whereby the emission (or absorption in the case of occultation measurements) of electromagnetic radiation from the atmosphere at the limb of the earth is observed from a satellite platform. The radiance $P$ observed by the instrument is given by the integral form of the radiative transfer equation

$$P = \int_t \int_\Omega \int_\nu A\left[\nu, \Omega(t)\right] \left\{ I_\infty(\nu, \Omega)\tau\left(\nu, \infty\right) + \int_{s=\infty}^{s=0} \tau\left(\nu, s\right) \frac{dB\left[\nu, T(s)\right]}{ds} ds \right\} d\nu \, d\Omega \, dt, \qquad (3.1)$$

where $\nu$ is frequency and $\Omega$ is solid angle, with $A\left(\nu, \Omega\right)$ describing the instrument's spectral and field of view response, and $\Omega(t)$ describing the movement of the MLS field of view as a function of time. $I_\infty(\nu, \Omega)$ represents the background emission, in this case the microwave background field. $s$ is the distance along a given ray path, where the spacecraft is at $s = 0$, $T(s)$ is the atmospheric temperature along this ray path. $B(\nu, T)$ is the blackbody function, which describes the thermal emission of the atmosphere as a function of temperature and frequency (for this discussion local thermodynamic equilibrium has been assumed, and the effects of scattering have been neglected.) The quantity $\tau$ describes the transmission of the atmosphere from the point $s$ to the spacecraft. This is defined by

$$\tau(\nu, s) = \exp\left[ - \int_{s'=s}^{s'=0} k\left(\nu, T(s'), f(s')\right) \rho(s') \, ds' \right]. \qquad (3.2)$$

The quantity $k$ is the absorption coefficient as a function of frequency, temperature and atmospheric composition described by the function $f(s)$. $\rho(s)$ is the atmospheric density.

## 3.2 Introduction to retrieval theory

Equation 3.1 gives an expression for the observed radiance as a function of the state of the atmosphere (i.e. its temperature and composition). The aim of the algorithms described in this document is to invert this calculation and obtain an estimate of the state of the atmosphere based on the observed radiances. Retrieval theory, the method by which these inverse calculations are constructed, has a great heritage in the remote sounding field (see Rodgers 1976; Rodgers 1990). The essential details of the subject, required for the MLS Level 2 algorithms, are given here.

### 3.2.1 The state vector, measurement vectors and $\chi^2$

It is clearly impossible to invert Equation 3.1 to obtain the functional form of the atmospheric temperature and composition profiles, as this would involve obtaining infinite degrees of freedom from a finite series of measurements. The problem can only be solved if a *state vector* is used. The state vector $\mathbf{x}$ is an $n$ element vector which describes all the aspects of the atmosphere and measurement system that affect the radiance measurements. Typically the state vector will contain profiles of temperature and composition represented by a finite set of vertical levels. By describing the state using a finite length vector, the retrieval task has been made tractable.

Measurements are also grouped into vectors, $\mathbf{y}_i$. Multiple vectors are used, as this explicitly indicates which sets of measurements are independent (i.e. have no covariance), and which are interdependent. For example, it is possible that radiometric noise may be correlated from channel to channel within an MLS spectral band, but not between radiometers. In such a case, one would use separate vectors to represent the radiances from each band. The covariance of the measurement vectors is represented by the matrix $\mathbf{S}_i$. In many cases, the measurement covariance matrices are purely diagonal. While it would be possible to thus split the corresponding measurement vectors up into separate one-element vectors, the grouping will typically be maintained for clarity. The computer programs that implement the retrieval algorithm will ensure that unnecessary calculations will be avoided in these cases, by considering only the diagonal elements of the matrix.

A key component of the retrieval algorithm is the *forward model*, this is a calculation which gives an estimate of the radiances that would be observed by the instrument, were the atmosphere in the state give by $\mathbf{x}$.

$$\hat{\mathbf{y}}_i = \mathbf{f}_i(\mathbf{x}), \tag{3.3}$$

where $\hat{\mathbf{y}}_i$ denotes an estimate of the vector $\mathbf{y}_i$. These functions are typically discrete forms of the radiative transfer integration in Equation 3.1. The aim of retrieval theory is to seek a value of the state vector which is a 'best estimate' of the true state of the atmosphere. The 'best estimate' is defined as that which most appropriately fits the observed measurements, by giving the minimum value of the quantity $\chi^2$, defined by

$$\chi^2 = \sum_i \left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x})\right]^{\mathrm{T}} \mathbf{S}_i^{-1} \left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x})\right]. \tag{3.4}$$

This expression is simply the vector form of the more familiar definition of $\chi^2$ in the scalar case as $\sum_i ([o_i - p_i]/s_i)^2$, where $o$ is the observed measurement, $p$ is the prediction, and $s$ is the standard deviation of the uncertainty in the observed data. In this case, the differences between the measured and fitted observations are weighted according to their covariances, and summed.

### 3.2.2 Newtonian iteration

Many different techniques exist for finding the minima of quantities such as $\chi^2$, the most commonly used method in retrieval algorithms is Newtonian iteration. The Newtonian minimization method gives an iterative expression for the best estimate of $\mathbf{x}$ according to

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} - \left[\nabla^2\chi^2\right]^{-1} \nabla\chi^2, \tag{3.5}$$

where $\nabla$ is the vector derivative operator

$$[\nabla]_i = \frac{\partial}{\partial x_i}, \tag{3.6}$$

and $\nabla^2$ is the matrix second derivative operator

$$\left[\nabla^2\right]_{ij} = \frac{\partial^2}{\partial x_i \partial x_j}. \tag{3.7}$$

Note that much of the retrieval theory literature refers to the retrieved estimate of the state vector as $\hat{\mathbf{x}}$, with $\mathbf{x}$ describing the unknown true state of the atmosphere. However, in this document, due to the large number of subscripts, superscripts and diacritics that $\mathbf{x}$ will soon gain, the 'hat' has been omitted for clarity. The exception is in Section 3.2.5 where the distinction between $\mathbf{x}$ and $\hat{\mathbf{x}}$ is important.

The parenthetical superscripts indicate values from different iterations. The vector calculus identities shown in Appendix C.1, when applied to Equation 3.4 give

$$\nabla\chi^2 = -2\sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x})\right] \tag{3.8}$$

$$\nabla^2\chi^2 = 2\sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{K}_i\,, \tag{3.9}$$

where

$$\mathbf{K}_i = \frac{\partial \mathbf{f}_i(\mathbf{x})}{\partial \mathbf{x}} \tag{3.10}$$

are known as the matrices of *weighting functions* (in some literature referred to as the *Jacobians*) for the measurement vectors. Note that we neglect second order terms here (i.e. those involving $\partial\mathbf{K}_i/\partial\mathbf{x}$), because we are in an iterative process, assuming linearity each iteration.

Thus Newtonian iteration can be expressed as

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} + \left[\sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1} \sum \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x}^{(r)})\right] \tag{3.11}$$

The covariance matrix of the solution can be shown to be given by

$$\mathbf{S}_{\mathbf{x}} = \left[\sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1} \tag{3.12}$$

### 3.2.3 The need for virtual measurements

In many cases the matrix inversion in Equation 3.11 is impossible as the matrix is singular. This indicates that the 'direct' measurements (radiances etc.) have provided insufficient information to completely determine the state vector; there are some components (or, more correctly, eigenvectors) of the system about which no information has been obtained.

The solution to this problem is to introduce *virtual measurements*. These are additional measurement vectors included in the retrieval calculation in order to assure successful matrix inversion, and to ensure reasonable values for comparatively poorly-measured aspects of the system. In the MLS case, as is typical, these virtual measurements take the form of *a priori* estimates of the state vector or individual components of the state vector, constructed from datasets such as climatologies. The covariance of the *a priori* information is chosen so as to limit the amount of bias in the resulting state vector.

Introducing *a priori* information in this manner can lead to incorrect interpretation of retrieved results. The covariance of the retrieved state vector should always be compared with the *a priori* covariance; if only a small amount of error reduction has been achieved, this indicates that the direct

measurements (i.e. radiances etc.) have failed to contribute significant information to the knowledge of the state vector.

In order to simplify later results, the *a priori* information is explicitly separated from the other measurement vectors. The *a priori* state vector is denoted by $\mathbf{a}$, with covariance given by the matrix $\mathbf{S_a}$. The forward model for this quantity is simply $\mathbf{f}_a(\mathbf{x}) = \mathbf{x}$ giving a corresponding $\mathbf{K}$ matrix equal to the identity. This gives a modified form of Equation 3.11 as

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} + \left[\mathbf{S_a}^{-1} + \sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1} \left\{\mathbf{S_a}^{-1}\left[\mathbf{a} - \mathbf{x}^{(r)}\right] + \sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x}^{(r)})\right]\right\}. \quad (3.13)$$

The solution covariance is in this case given by

$$\mathbf{S_x} = \left[\mathbf{S_a}^{-1} + \sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1} \quad (3.14)$$

These equations form the basis of all the MLS retrieval calculations.

In some retrieval situations there are elements of $\mathbf{x}$ for which the use of an *a priori* as a virtual measurement is inappropriate, for these elements, the corresponding rows and columns of $\mathbf{S_a}^{-1}$ are set to zero [1].

### 3.2.4 Retrieval phases and errors on constrained quantities

In many cases, the retrieval calculations are performed in separate phases, with the results from one phase being used in the forward model calculations for later phases. For example, in the UARS MLS data processing algorithms, a retrieval of temperature and tangent point pressure was obtained from the 63 GHz $O_2$ radiances. These results were then used in retrievals of the constituent information from the other spectral bands.

However, when performing a retrieval calculation in separate phases, the measurement covariance matrices $\mathbf{S}_i$ should be modified to allow for the fact that there are uncertainties in the knowledge of the previously-retrieved quantities. The modification should be made according to

$$\mathbf{S}_i \rightarrow \mathbf{S}_i + \mathbf{K}_{i[\mathbf{c}]}\mathbf{S_c}\mathbf{K}_{i[\mathbf{c}]}^{\mathrm{T}}, \quad (3.15)$$

where $\mathbf{S_c}$ describes the covariance of the quantities $\mathbf{c}$ that were previously retrieved (i.e. the covariances obtained from Equation 3.14,) and the matrices $\mathbf{K}_{i[\mathbf{c}]}$ are the weighting functions for these quantities, such that

$$\left[\mathbf{K}_{i[\mathbf{c}]}\right]_{\alpha j} = \frac{\partial\,(\mathbf{y}_i)_\alpha}{\partial \mathbf{c}_j}. \quad (3.16)$$

This calculation is sometimes referred to as *constrained quantity error propagation*.

As described in Section 3.2.1, in many retrieval problems the measurement covariance matrices are diagonal (or can be assumed to be diagonal to a reasonable level of approximation.) This is the case for most of the MLS spectral bands (see the discussion in Sections 3.5.3 and 3.5.4). If a diagonal covariance matrix can be assumed, then computation time can be saved by optimizing the algorithm to take advantage of this fact. However, a constrained quantity error propagation calculation will typically produce a new set of measurement covariance matrices that are not diagonal. As these measurement covariance

[1] Note that strictly speaking, this makes $\mathbf{S_a}^{-1}$ singular. A more detailed discussion of the construction of the *a priori* state vector and its covariance matrix is given in Section 4.2.3.

matrices have to be inverted as part of Equations 3.13 and 3.14, this can represent a significant amount of computational effort. In particular, in the case where the number of measurements greatly exceeds the size of the state vector (as in the MLS case), it is generally preferable to avoid constrained quantity error propagation. Instead, the most efficient approach is to retrieve all the elements of the state vector simultaneously, using all the measurement vectors. In this manner, the measurement covariance matrices $\mathbf{S}_i$ remain diagonal. This does not preclude the use of phasing, however. Phase can be implemented in a different manner, such that more quantities are added to the state vector each phase, rather than considering a completely different set of quantities each phase. Sections 4.4.2 and 7.1 deal with these issues in more detail.

### 3.2.5 Diagnosing retrieval performance

#### Averaging Kernels

When examining the results of a retrieval calculation it is important to check the retrieved error estimate and compare them with any *a priori* information as outlined in Section 3.2.3. One way in which to do this comparison is to look at the *averaging kernel* matrix. This is defined as

$$\mathbf{A} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}} = \left[ \mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^\mathrm{T} \mathbf{S}_i^{-1} \mathbf{K}_i \right]^{-1} \sum_i \mathbf{K}_i^\mathrm{T} \mathbf{S}_i^{-1} \mathbf{K}_i, \tag{3.17}$$

where for the sake of this discussion, $\mathbf{x}$ is the true state vector, with $\hat{\mathbf{x}}$ as the retrieved state from Equation 3.14. In this calculation, as well as distinguishing between the true and retrieved state vectors $\mathbf{x}$ and $\hat{\mathbf{x}}$, it is necessary to distinguish between the *forward model* $\hat{\mathbf{y}}_i = \hat{\mathbf{f}}_i(\hat{\mathbf{x}})$, which provides an estimate of the observed radiance given an estimate of the state vector, and the *forward function* $\mathbf{y}_i = \mathbf{f}_i(\mathbf{x})$, which describes the actual physical process taking place in the atmosphere, aspects of which (e.g. spectroscopic constants) are uncertain.

$\mathbf{A}$ describes the sensitivity of the retrieval to the true atmospheric state, as opposed to its sensitivity to the *a priori*. One could consider it as the 'ratio' of the information contributed by the direct measurements compared to the total contributed by both the direct measurements and the *a priori*.

Columns of the $\mathbf{A}$ matrix represent the response of the retrieval system to a 'delta function' disturbance in the atmosphere (i.e. a change in a single element of $\mathbf{x}$.) Rows of the matrix indicate the amount each element of the true state vector has contributed to the retrieved estimate.

#### The use of $\chi^2$ as a diagnostic

In addition to examining these quantities, the $\chi^2$ value for a retrieval (or for each measurement vector independently) should also be examined. Ideally, the value of $\chi^2$ should be about $m$, where $m$ is the number of elements in the measurement vector(s) under consideration ($\chi^2$ is sometimes divided by $m$ to yield a 'normalized' $\chi^2$ value.)

A value of $\chi^2$ significantly larger than $m$ indicates that the radiance measurements have not been fitted to a sufficient level of accuracy. This can be due to errors in the forward model, or poor convergence in the retrieval algorithm. A $\chi^2$ significantly lower than $m$ on the other hand usually indicates that the measurement precisions used are too pessimistic.

### 3.2.6 Other minimization techniques

In many cases, (e.g. for some of the MLS observations) the retrieval calculation is sufficiently linear that a small number of Newtonian iterations can yield the correct result. However, in some cases, such as

those where the system is moderately non-linear, and the initial value of the state vector is comparatively far from the solution, other techniques may be more appropriate.

The steepest descent approach is a simple algorithm which makes small steps each iteration in the direction of the steepest descent of the cost function ($\chi^2$ in this case.)

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} - \gamma^{-1}\nabla\chi^2, \tag{3.18}$$

where $\gamma$ is a scalar value describing the size of step to be taken. This is typically a slow algorithm, as it takes no advantage of possible linearity in the system.

The Marquardt-Levenberg approach is faster, as it is a combination of the 'cautious' steepest descent method and the 'aggressive' Newtonian method. As the iterations proceed, and the solution is approached, the steps taken each iteration become smaller, making linearity an increasingly better assumption, and allowing the minimization to become more aggressive. The iterations proceed according to

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} - \left[\gamma\mathbf{I} + \nabla^2\chi^2\right]^{-1}\nabla\chi^2. \tag{3.19}$$

For small values of $\gamma$ this is equivalent to the Newtonian iteration in Equation 3.5, while for large values of $\gamma$ this is equivalent to steepest descent with a small step size. As the iterations proceed the value of $\gamma$ is changed according to the convergence behavior. If $\chi^2$ increases, the new value of $\mathbf{x}$ is rejected, and $\gamma$ is increased so that a more cautious step can be explored. If $\chi^2$ decreases, then the new value of $\mathbf{x}$ is adopted and $\gamma$ is decreased so that the next iteration is more aggressive. The use of the Marquardt-Levenberg method can lead to more stable convergence in most moderately non-linear cases than Newtonian iteration. For more information on the Marquardt-Levenberg scheme, see Press et al. 1986; for applications to retrieval theory see Marks and Rodgers 1993.

## 3.3 Radiance measurements

The MLS instrument makes observations of microwave radiation in many different regions of the spectrum, covering the frequency range from 118 GHz to 2.5 THz. The instrument consists of seven microwave radiometers, covering five different spectral regions. The signals from the radiometers are passed onto various spectrometers. The spectral coverage of the instrument is shown in Figure 3.1. In the MLS data processing, radiance is measured in Kelvins and considered to be a *brightness temperature*. This is a quantity proportional to the observed radiance, which, in the long wavelength limit, is equal to the temperature of the blackbody producing the observed radiance.

In the standard operational mode, the instrument makes one complete vertical scan of the GHz antenna over tangent heights between 2.5 and 62.5 km in 20 s. The THz antenna scans from 15.0 to 62.5 km in the same time period. Approximately 4.7 s are spent in calibration and antennae retrace activities, giving a repeat period of about 24.7 s. Each scan/calibrate/retrace activity is called a *major frame*. The lengths of the major frames will vary slightly in increments of 1/6 s to ensure that the latitudinal distribution of the scans is the same from orbit to orbit, and between both ascending and descending orbital segments, and the northern and southern hemispheres. The 20 s limb scan is a continuous movement, as opposed to the 'stop and stare' scan that was used in UARS MLS. During the scan, 120 radiance integrations are performed, each of length 1/6 s. These integration periods are known as *minor frames*.

### 3.3.1 Behavior of the radiances

Figure 3.2 shows a set of calculated radiances from two of the EOS MLS spectral bands. The form of the radiance curves shown in the left hand plot is typical of the observations from limb sounding

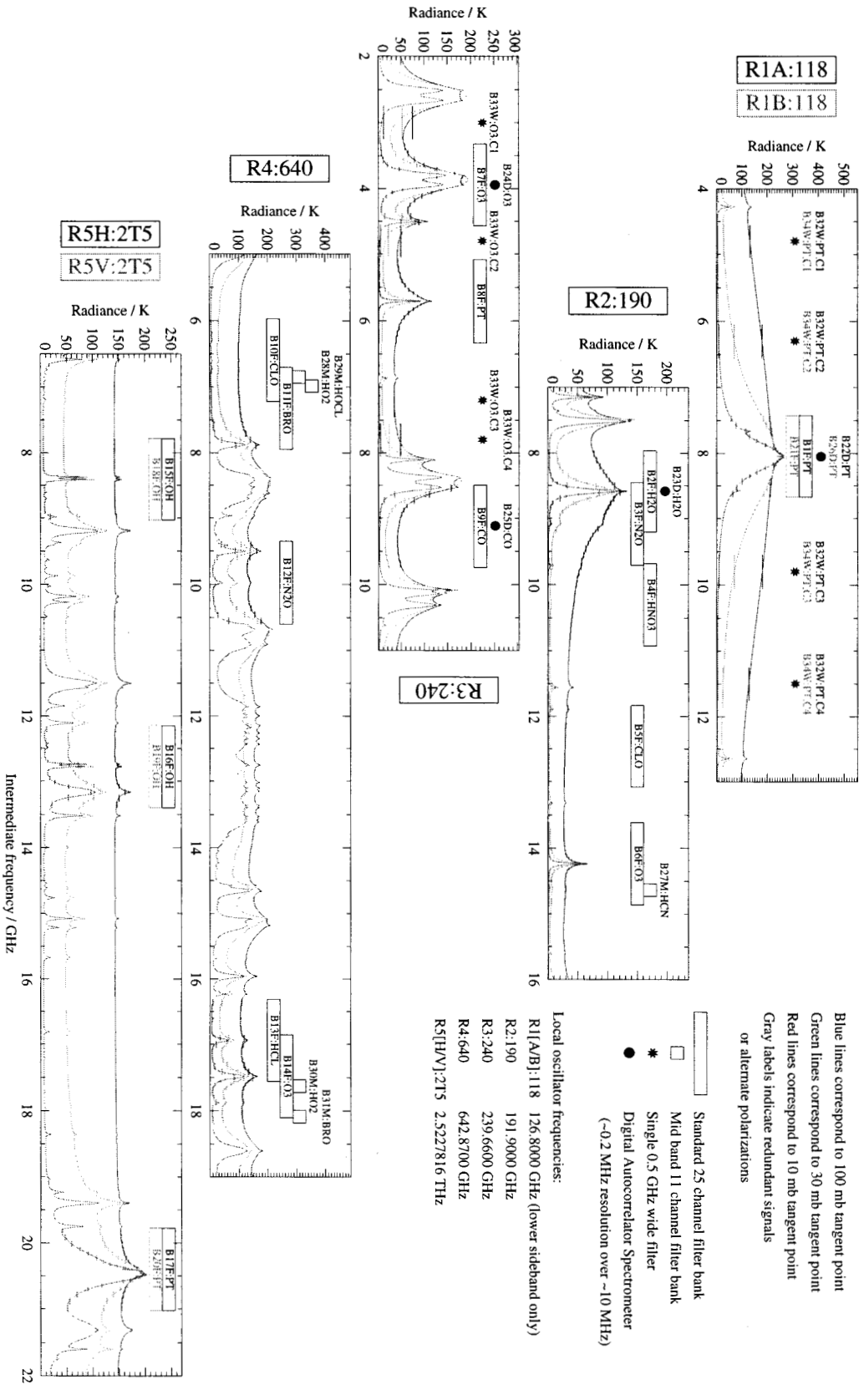
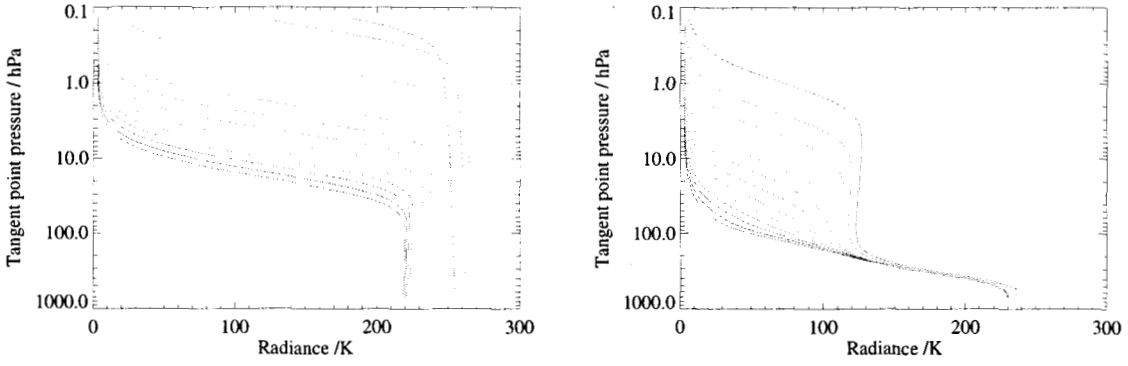



**Figure 3.1:** (color) EOS MLS Spectral coverage. See Appendix E for a description of the EOS MLS signal designation nomenclature.




**(a)** `R1A:118.B1F:PT`.

**(b)** `R2:190.B2F:H2O`.

**Figure 3.2:** (color) These plots show example radiance profiles for two different spectral bands in EOS MLS. The `R1A:118.B1F:PT` band is targeted at an $O_2$ line, and used to measure temperature and tangent pressure. The `R2:190.B2F:H2O` band is used to measure $H_2O$. Only the radiances for the first 13 channels in each band have been shown. The red lines correspond to the channels closest to the centers of the spectral lines, with the purple lines corresponding to those on the line wings. See Appendix E for a description of the EOS MLS signal designation nomenclature.

instruments. At high tangent point altitudes, the atmospheric density along the limb path is very low, so little emission is observed. As the ray path descends through the atmosphere, emission becomes stronger as the atmosphere becomes thicker. Eventually, the atmosphere becomes sufficiently opaque that emission from lower regions in the atmosphere is absorbed by air at higher altitudes, and is thus never observed by the instrument. In these circumstances the radiances are said to be *saturated* or *blacked out*, this is the cause of the knee in the radiance curves. The saturated radiances, are a measure of the temperature of the region of the atmosphere where the saturation takes place. Sometimes, the radiances continue to increase or decrease slightly as the tangent ray path is scanned further down. This is a geometrical effect. As the ray descends, the path length to a given height decreases, thus the saturation occurs at a different height in the atmosphere, leading to slightly different radiances, depending on the form of the temperature profile. The channels closer to the line centers will saturate at higher altitudes than those on the wings of the lines, as the absorption at the frequencies closer to the line centers is stronger.

The curves shown in the right hand plot in Figure 3.2, corresponding to the 183 GHz $H_2O$ observations, show a slightly different behavior. In this case two separate saturation process occur. This is simply explained by the fact that these observations are made by a 'double sideband' radiometer. The MLS radiometers work on the microwave heterodyne technique; they output an *intermediate frequency* (IF) signal, corresponding to the observed signal with the *local oscillator* (LO) frequency subtracted. The negative frequency components are folded over into positive IF space with a 180° phase shift. This is described in Figure 3.3. The intermediate frequency signal is given by

$$\langle \textit{Intermediate frequency signal} \rangle = \alpha \, \langle \textit{Upper sideband signal} \rangle + (1 - \alpha) \, \langle \textit{Lower sideband signal} \rangle \, . \quad (3.20)$$

The *sideband ratio*, $\alpha / (1 - \alpha)$, is measured during prelaunch calibration.

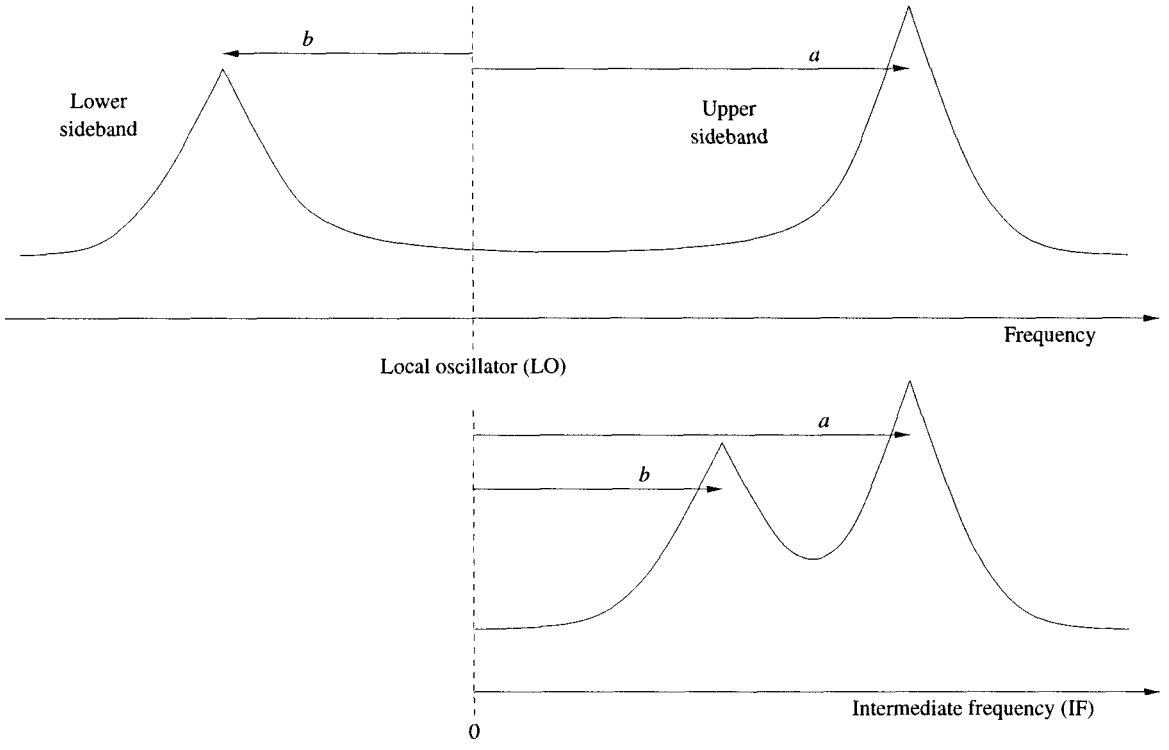



**Figure 3.3:** This figure shows the action of the microwave heterodyne technique. The observed radiances are combined in a non-linear element such as a diode, with a *Local Oscillator* (LO) signal. The resulting *Intermediate Frequency* (IF) signal is a combination of the signals in the upper and lower sidebands.

### 3.3.2 The importance of tangent pressure

As shown in Figure 3.2, the radiances observed by MLS depend strongly on the atmospheric pressure at the ray tangent point. This is due both to the large increases in atmospheric density with decreasing altitude, and to the fact that the spectral lines being measured are, at most altitudes, pressure broadened (an effect caused by collisions between emitting molecules). At higher altitudes (in the mid to upper mesosphere), pressure broadening becomes less significant, and Doppler broadening due to the distribution of molecular velocity takes over as the dominant effect.

Quantities such as radiance and tangent pressure are somewhat distinct from explicitly geophysical quantities, such as temperature and composition, in that they are to a greater or lesser extent dependent on the state of the instrument (e.g. the pointing of the antenna.) These are often referred to as *minor frame quantities*, in that they vary from one MLS minor frame to the next. Where it is useful to draw distinctions between such quantities and strictly geophysical parameters, an arrow will be drawn over the relevant symbol, thus the tangent point pressure for minor frame $i$ would be represented by the symbol $\vec{p}_i$.

## 3.4 Geometric measurements

In addition to microwave radiances, the MLS instrument's knowledge of the altitude of the limb tangent point $\vec{h}_i$, can be considered as a measurement. Given knowledge of the tangent point pressures $\vec{\mathbf{p}}$, the atmospheric temperature profile as a function of pressure $T(p)$, and the altitude $h_0$ of a fixed pressure

surface $p_0$, a hydrostatic integration will serve as a forward model for such measurements.

$$\vec{h}_i = h_0 - \int_{p_0}^{\vec{p}_i} \frac{R(p)T(p)}{g(p)p} dp, \tag{3.21}$$

where $g$ is the acceleration due to gravity and $R$ is the atmospheric gas 'constant'.

The use of the additional information provided by these measurements is essential in supplying tangent pressure information in regions where the radiances are largely independent of pressure such as the lower altitude regions where the radiances are saturated.

The problem with this construction of the forward model is that the quantity $g$ is not a function of pressure, but of geometric altitude. One could express $g$ as a function of pressure, however a simpler approach is to recast this expression as one for the geopotential height $\vec{Z}$ of the tangent point, defined as

$$Z = Z_0 + \frac{1}{g_0} \int_{h_0}^{h} g \, dh, \tag{3.22}$$

where $g_0$ is a nominal value of $g$ (usually taken as 9.80665 $ms^{-2}$.) When expressed in terms of geopotential height, Equation 3.21 becomes

$$\vec{Z}_i = Z_0 - \int_{p_0}^{\vec{p}_i} \frac{R(p)T(p)}{g_0 p} dp. \tag{3.23}$$

Equation 3.23 can thus form the basis of a forward model if the Level 1B tangent point altitude is converted into geopotential height (a well understood and documented calculation). These geopotential heights are then used as the measurements, rather than the geometric altitudes.

There is however an additional complication, due to the refraction of the tangent ray by the atmosphere, this leads to a difference between the true ray tangent point altitude (and thus geopotential height) and the 'unrefracted' tangent point altitude as determined by geometry. The magnitude of this refraction is a function of atmospheric density (and thus temperature and pressure), and humidity.

This effect leads to a conceptual problem with the measurement system, in that the measurements depend on the value of the state vector, which is not strictly speaking valid in retrieval theory (only the predicted measurements obtained from the forward model should depend on the value of the state vector.) This is really only a matter of semantics, the situation can easily be remedied by defining the measurement as a *scan residual*, the difference between the hydrostatic and geometric/refraction calculations of tangent point geopotential height, defining the values of this measurement to be zero. The forward model will be constructed to compute this difference. The covariance matrix for this measurement will be constructed as a diagonal matrix, with the diagonal values being based on the precision reported on the Level 1B tangent point altitude data, and no off diagonal correlations.

## 3.5 Construction of the MLS 'state vector'

Section 3.2 gave an introduction to the concept of the 'state vector', the vector **x** that describes all the aspects of the atmosphere and measurement system that can affect the direct measurements. The construction of the state vector is an essential part of the design of any retrieval system; its contents need to be chosen carefully to provide a complete set of independent parameters that describe the whole system. This section discusses the construction of the state vector for the MLS retrieval system.

### 3.5.1 State vector selection methodology and implementation

The construction of the MLS state vector is performed by studying the measurement system with reference to a set of simple criteria.

Firstly, the experimental objective is placed in the state vector. That is profiles of atmospheric temperature **T**, geopotential height **Z** and composition (concentration of targeted gases) **f** on fixed pressure surfaces.

At this point, any superfluous information in the state vector is removed. In this case the temperature information is redundant with most of the geopotential height profile as, given profiles of temperature and one geopotential height, the entire **Z** profile could be computed using hydrostatic balance. For this reason, the vector is reduced so as to contain only one geopotential height.

Next, the primary source for direct information about the state vector is identified. In the MLS case, this direct information is the radiance observations $\vec{\mathbf{P}}$ and the scan residual measurements $\vec{\mathbf{r}}$.

The next set of quantities to be placed in the state vector represents any additional information needed in order to characterize the direct measurements. In this case, it is clear that tangent pressure information for each radiometer is essential if forward model estimates are required for both radiances and scan residual measurements. In order that the calculations may be more linear, this quantity is represented by the vector $\vec{\zeta} = -\log_{10}[\text{Tangent pressure / hPa}]$.

The use of tangent pressure is slightly complicated by the fact that the MLS instrument performs a continuous scan. The quantity stored in the state vector represents the tangent point pressure at some prescribed time within the integration period. The forward model process will perform the appropriate interpolations and convolutions to account for the movement of the tangent point during the integration.

This whole process is somewhat iterative, in that these new quantities may themselves be dependent on further additional information. For example, the estimates of tangent pressure for each radiometer can be dramatically improved given knowledge of the angular offsets between the various radiometers and/or modules in the case of the GHz/THz module offsets. Thus these offsets are included in the state vector. These offsets are essentially constant, although in the case of the offset between the GHz and THz modules, there may be a slight orbital dependence. Additional redundancy may develop as more quantities are introduced into the state vector. It is important to recognize the source of this redundancy and attempt to eliminate it.

A pictorial representation of this whole process, as applied to the EOS MLS measurement system is given in Figure 3.4.

As the state vector is constructed, attention needs to be paid to the possible need for *a priori* values for the elements added. In the case of the $\vec{\zeta}$ components, any *a priori* values would need to be based on the observed Level 1B tangent point height information. Clearly, as these heights are being used (in the form of geopotential heights) as direct measurements as described in Section 3.4, the use of an *a priori* for $\vec{\zeta}$ would result in the use of the same information twice. In fact, an *a priori* estimate for $\vec{\zeta}$ is unnecessary, as the measurement system already has enough information to completely describe $\vec{\zeta}$. The Level 1B tangent point height, and the temperature elements of the state vector provide enough information to describe values of $\vec{\zeta}$ even in cases where all the radiances for a given minor frame are missing.

### 3.5.2 Representation within the state vector

As described in Section 3.2.1, the state vector is designed to represent the functional form of the atmospheric temperature and composition. In the MLS case, this functional form is constructed using a set of *basis functions*. For example, in the case of the atmospheric temperature profile, the state vector is

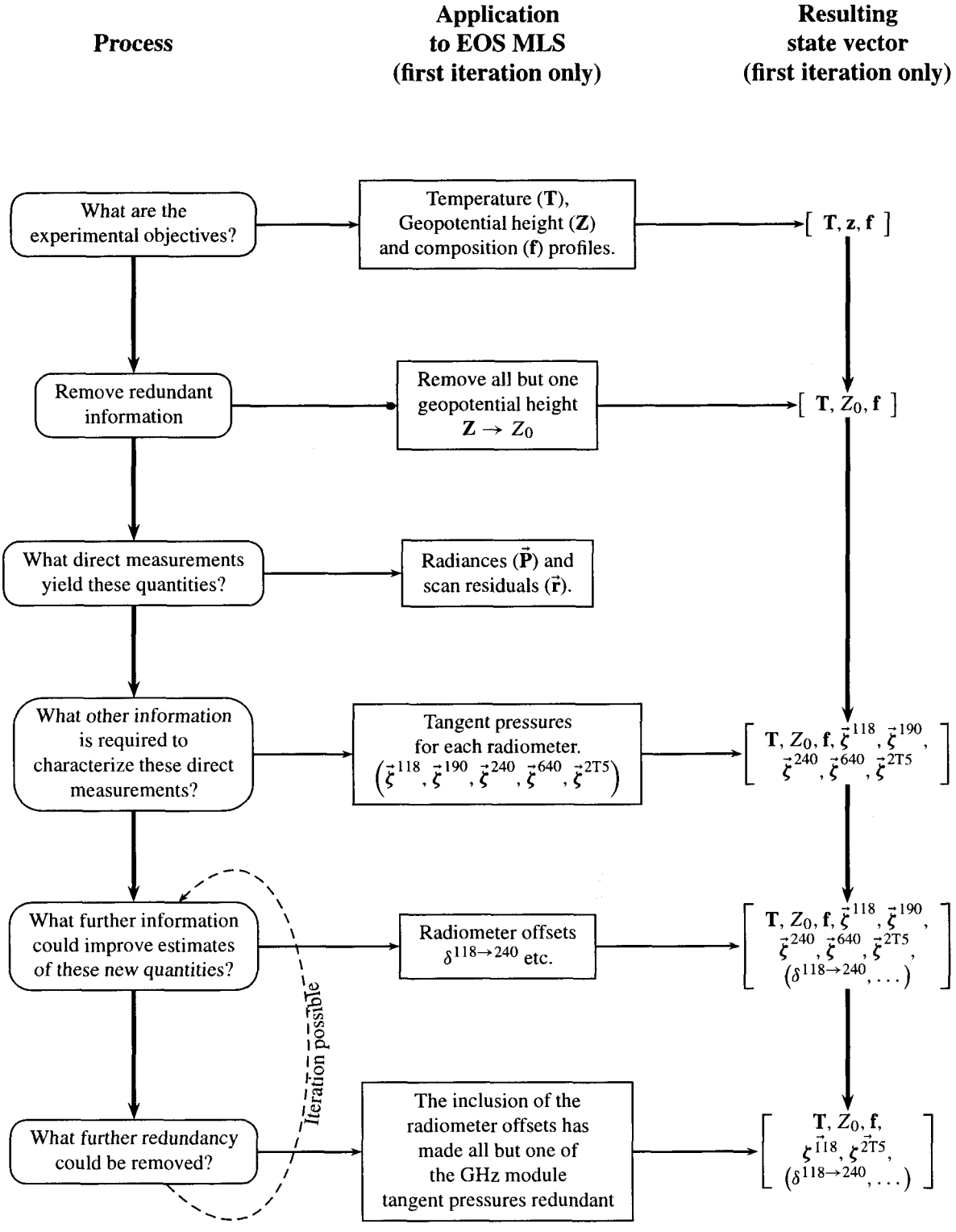



**Figure 3.4:** A pictorial representation of the EOS MLS state vector selection process

defined according to

$$T(z) = \sum_i T_i \eta_i(z), \tag{3.24}$$

where $z$ is the log pressure vertical coordinate[2] and $T(z)$ is the functional form of temperature. $T_i$ are the components of the state vector describing the temperature, and $\eta_i(z)$ are the basis functions.

For most of the MLS species, the $n$ basis functions $\eta_0 \ldots \eta_{n-1}$ are defined as

$$\eta_i(z) = \begin{cases} 1 & \text{if } (i = 0 \text{ and } z < z_0) \text{ or } (i = n-1 \text{ and } z > z_{n-1}), \\ \frac{z_{i+1}-z}{z_{i+1}-z_i} & \text{if } z_i \le z < z_{i+1}, \\ \frac{z-z_{i-1}}{z_i-z_{i-1}} & \text{if } z_{i-1} \le z < z_i, \\ 0 & \text{otherwise} \end{cases} \tag{3.25}$$

A sample set of such functions are shown in Figure 3.5. These functions give a profile whose functional form is equivalent to linear interpolation in $z$ between the values $T_i$ on surfaces $z_i$, with no extrapolation beyond $z_0$ or $z_{n-1}$.

For some species, this linear interpolation form may not describe the true atmospheric profile with sufficient accuracy. An example of this is the case of tropospheric $H_2O$ concentration. The vertical profile of $H_2O$ in the upper troposphere shows rapidly decreasing abundance with increasing altitude. In order to accurately capture this feature, alternative representation bases may be appropriate. In the case of UARS MLS, an exponential form of basis was used to capture these features. This was achieved by retrieving a state vector describing a linear basis in relative humidity, which is equivalent to an exponential one in mixing ratio space. A similar approach will be implemented for EOS MLS, though the intended increase in the state vector resolution from 6 to 12 surfaces per pressure decade will also improve the accuracy of the representation.

### 3.5.3 Continuum emission and 'baseline'

Most of the MLS observations rely on measurements of *spectral contrast*, that is, the retrieval algorithms determine the atmospheric composition by effectively comparing the radiances near the center of a spectral line with those in the wings. The absolute value of the radiances observed is not typically as important a factor in the MLS measurement system. There are many factors that determine such spectrally flat contributions to the MLS radiance observations. The term 'flat' in this context implies signals that may vary between the spectral bands, but will vary only slowly within the bands. Unexplained spectrally flat signals can arise in the measurement system from many different sources, these broadly divide into instrumental and forward model factors.

Instrumental contributions include unmodeled blackbody emission from the MLS antenna, and uncertainties in the knowledge of the instrumental field of view. In the case of the field of view, an important factor is the uncertainty in the signal received through sidelobes of the antenna pattern.

The main mechanism whereby spectrally flat errors can be introduced into the forward model calculation is emission, absorption or scattering by species whose abundance is not known to a high enough accuracy (e.g. emission from cirrus ice in the upper troposphere), and/or whose emission is not well known at the frequencies under consideration (for example water vapor in the lower stratosphere / upper troposphere.) While the water vapor profile from about 500 hPa upwards will be measured by MLS,

[2]Do not confuse $z$ with $\vec{\zeta}$, $z$ is a fixed vertical coordinate, $\vec{\zeta}$ is a set of state vector elements describing the pressure at the limb path tangent points.

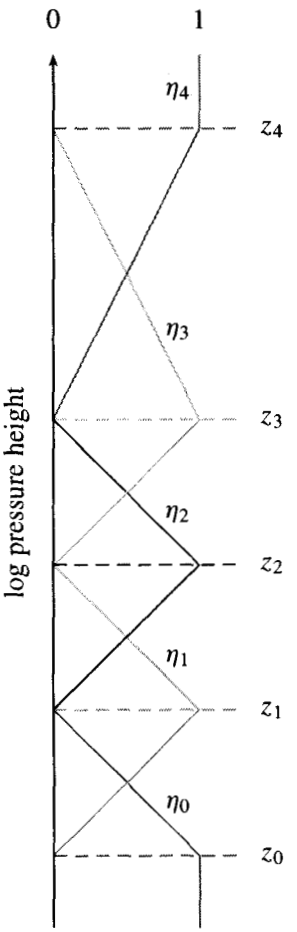



**Figure 3.5:** (color) This figure shows an example of the representation bases used in the MLS state vector. The functions $\eta_i$ are shown by the colored lines, corresponding to log pressure heights $z_i$. The vertical resolution need not be constant, and in this case, the resolution halves for the layer $z_3 \rightarrow z_4$.

the spectroscopy of the $H_2O$ emission is not sufficiently well known to account for the spectrally 'flat' radiance at the lower altitudes with an accuracy that is consistent with the MLS measurement accuracy.

It is possible to model these errors through the use of off diagonal terms in the MLS radiance measurement covariance matrices. However, as explained in Section 3.2.4, the computational effort required when dealing with such covariance matrices is vast. A more practical approach is to include elements in the state vector that model such errors, and retrieve them. This is the solution adopted in the MLS processing.

The instrumental contributions are typically best characterized in terms of an absolute radiance offset term in the forward model, while the forward model effects are best accounted for by retrieving a spectrally flat extinction coefficient on fixed pressure surfaces. Clearly, these two quantities are highly correlated, and retrieving an extinction profile and a minor frame dependent radiance offset would lead to a highly degenerate system. The approach adopted will be to retrieve the radiance offset on a once per major frame basis (or perhaps coarser) with the extinction only retrieved below some altitude, (e.g. 10.0 hPa), above which it is expected to be small. This should allow good separation between the instrumental and the forward model contributions.

### 3.5.4 Other sources of correlated radiance error.

In addition to the 'baseline' terms discussed above, there are additional mechanisms which can produce correlations in the errors on MLS radiance measurements. One such is the so-called $1/f$ noise described in Jarnot 1999. In most cases, this effect is well modeled by the retrieval of the baseline terms described above. One possible exception may be cases where the spectral signal being sought is very weak. These are the measurements that will be used to derive the 'noisy' products. For a full description of the issues involved with these measurements, see Chapter 5.

Several of the spectral bands in the MLS instrument overlap, as shown in Figure 3.1, for example `R2:190.B2F:H2O` and `R2:190.B3F:N2O`. The noise on the MLS radiances is dominated by noise from the radiometers, rather than noise from the individual spectrometers. This means that there will be a correlation in the errors on individual radiances from these two bands. If this is not modeled correctly, the retrieval algorithm will draw false inferences, believing it has two independent measures of the atmosphere, when in fact their noises are identical.

A full, non-diagonal, treatment of such cases is impractical, as described above. However, there are two possible alternative solutions. One could approximate the covariance matrix by increasing all the diagonal elements by a factor of $\sqrt{2}$. This prevents the retrieval algorithm reading too much into the measurement system, but is an approximation, as it does not indicate the coupled nature of the problem. The prefered approach is simply to only use data from one of the two channels in regions of overlap. Typically one would choose to keep the channels with higher frequency resolution. There would probably remain either small regions of overlap, or small gaps, due to the lack of complete alignment in the channels for the two bands, but these are of little consequence.

The spectral response of the MLS signal channels is not completely rectangular as it would be in the ideal case. This leads to a very slight overlapping of spectral response between adjacent channels. While the detailed spectral response of the channels will be considered fully in the forward model calculations, the small level of correlated uncertainty the overlap produces can safely be neglected as the overlap is very small.

### 3.5.5 A note on 'state vectors' and 'retrieval vectors'

So far the discussion in this section has assumed that all of the elements of the state vector will be retrieved by the MLS data processing algorithms. In reality, as described in Section 3.2.4, the retrieval

algorithm will proceed in separate phases. For example a retrieval of temperature and pressure may be followed by a retrieval of tropospheric humidity and ozone abundance, followed in turn by a retrieval of stratospheric composition. Thus, as the retrieval progresses, some elements of the state vector may be constrained to values, either taken from a prior phase, or from some *a priori* information, as described in Section 3.2.4.

In the context of the standard literature (e.g., Rodgers 1976; Rodgers 1990), the 'retrieved' and 'constrained' quantities are distinguished by defining the forward model as

$$\hat{\mathbf{y}} = \mathbf{f}(\mathbf{x}, \mathbf{b}), \tag{3.26}$$

where the vector **b** describes the quantities that are constrained. In this notation, the 'state vector' described in this section would be the combined vector [**x**, **b**] (some times called the 'forward model' vector), while the 'retrieval vector' would be the vector **x** alone.

In most of this document the distinction is unnecessary, so the term 'state vector' will be used to loosely refer to either the 'retrieval vector' or the 'forward model' vector.

### 3.5.6 The proposed state vector for MLS

The full details of the MLS state vector remain to be clarified. For a summary of the current proposed configuration see Appendix B.

## 3.6 Observation geometry

The EOS MLS instrument observes radiances emitted from the limb of the earth in the plane described by the spacecraft orbit. This contrasts with the UARS MLS instrument which observed limb emission perpendicular to the flight direction. This observation geometry has far reaching implications for improving the retrieval algorithms, best explained with reference to Figure 3.6. This figure makes it clear that the radiances observed in an individual limb scan are a function of the temperature and composition of the atmosphere over a range corresponding to several adjacent Level 2 profiles. In the UARS MLS case, while the horizontal scales involved were similar, spherical symmetry was assumed, as there was no direct way to obtain information about the horizontal variability of the atmosphere along the line of sight. Some instruments such as the *Improved Stratospheric And Mesospheric Sounder* instrument on UARS (Taylor et al. 1993), which had the same measurement geometry as UARS MLS, adopted a two pass approach, whereby the profiles obtained from a first pass retrieval were mapped onto a grid. The horizontal gradients of this gridded field were then used as *a priori* information in a second pass of the retrieval process (Dudhia and Livesey 1995).

In the EOS MLS case however, with the instrument looking along the flight direction, much more information on the horizontal variability of the atmosphere along the line of sight can be obtained. Spherical symmetry need no longer be assumed, and the retrieval task can be structured to account for, and provide information on, horizontal variability along the line of sight. The issues raised by this are covered in detail in the next chapter.

### 3.6.1 'Fundamental' coordinates

In constructing the state vector, it is very important to recognize which sets of quantities are independent, and which are related. For example, many of the aspects of the MLS measurement system are functions of earth radius and the acceleration due to gravity $g$ for a given altitude. However, both of these quantities are themselves functions of latitude (and to a lesser extent, longitude.) It is important

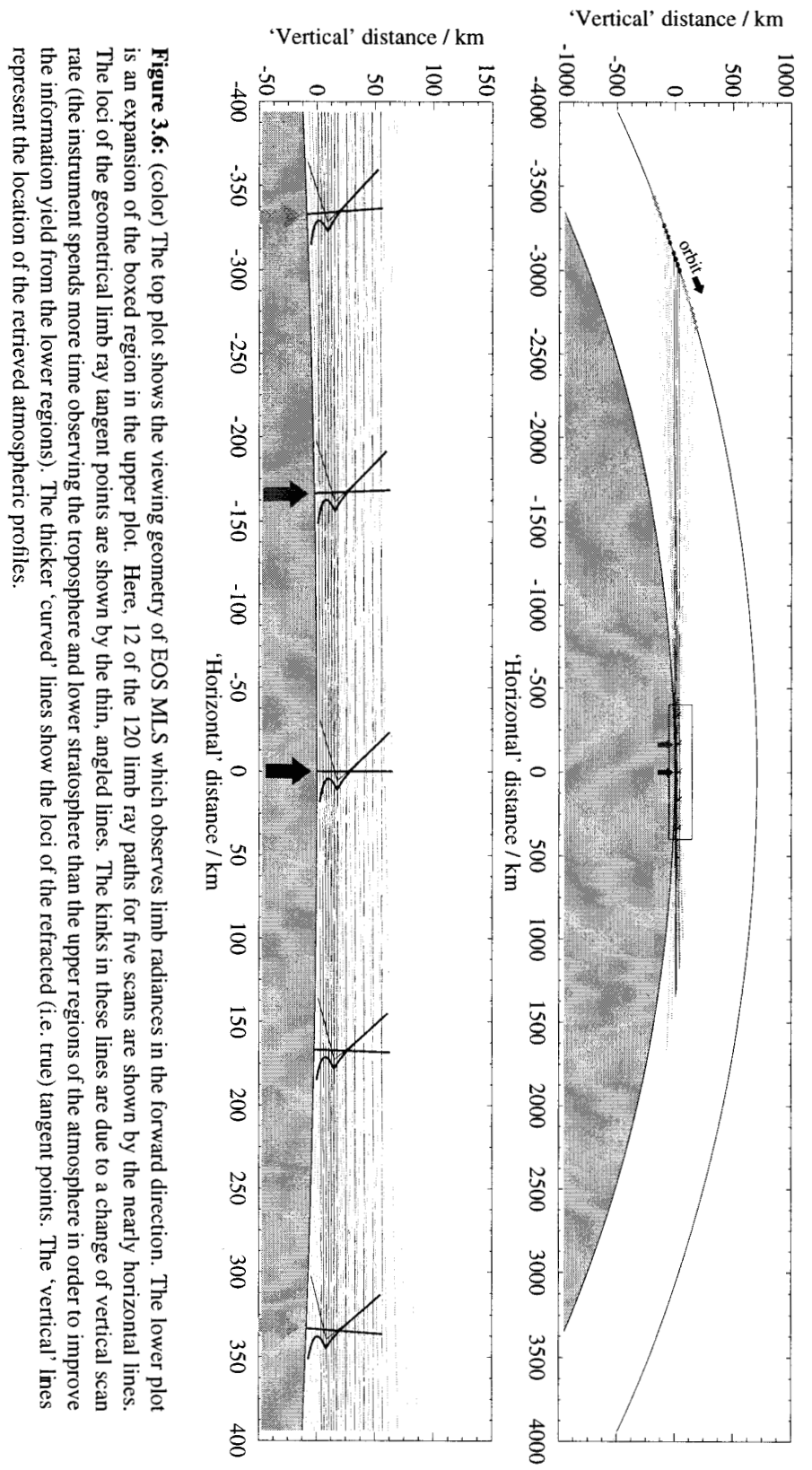



**Figure 3.6:** (color) The top plot shows the viewing geometry of EOS MLS which observes limb radiances in the forward direction. The lower plot is an expansion of the boxed region in the upper plot. Here, 12 of the 120 limb ray paths for five scans are shown by the nearly horizontal lines. The loci of the geometrical limb ray tangent points are shown by the thin, angled lines. The kinks in these lines are due to a change of vertical scan rate (the instrument spends more time observing the troposphere and lower stratosphere than the upper regions of the atmosphere in order to improve the information yield from the lower regions). The thicker 'curved' lines show the loci of the refracted (i.e. true) tangent points. The 'vertical' lines represent the location of the retrieved atmospheric profiles.

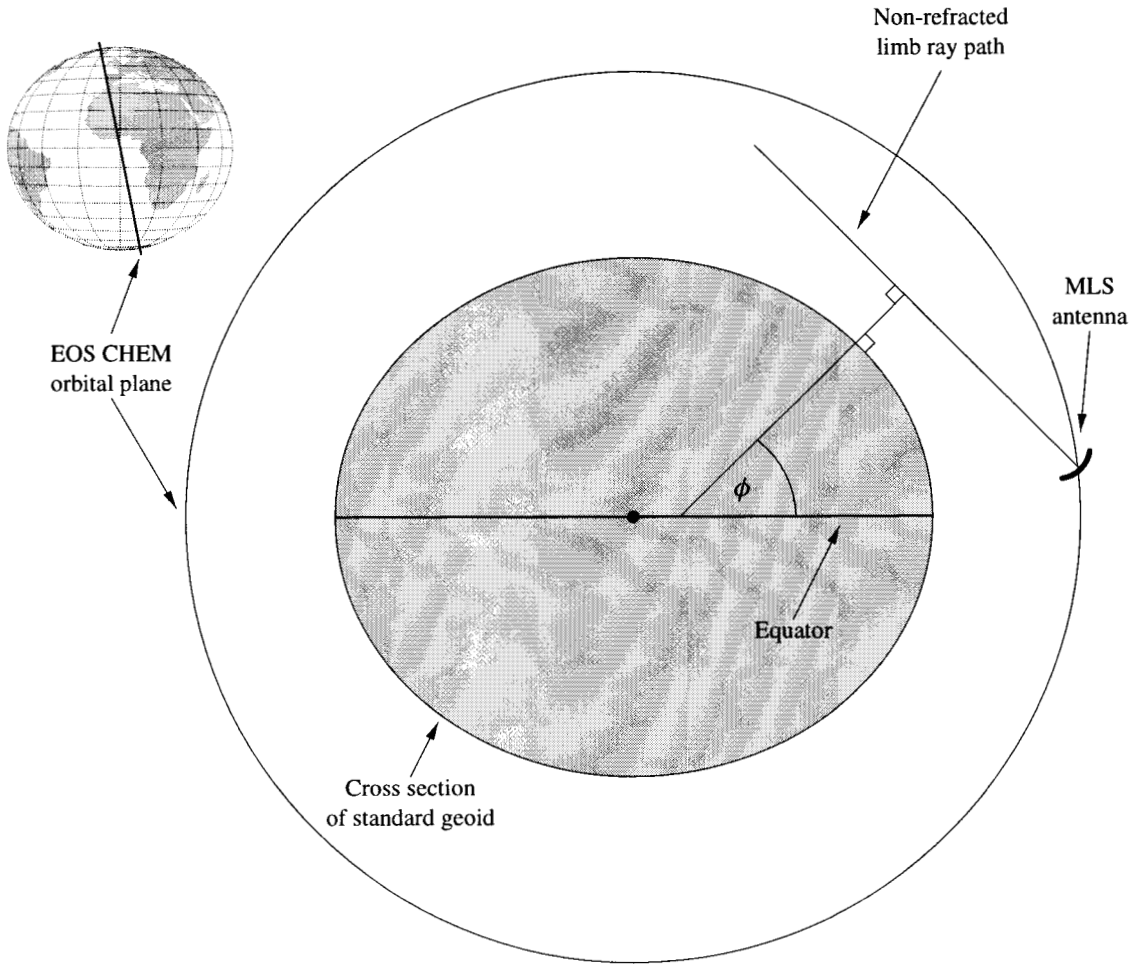



**Figure 3.7:** The observation geometry of MLS is affected by the oblateness of the earth, and the inclination of the orbit. The main part of this figure shows a cross section in the 98° inclined orbital plane. The master horizontal coordinate $\phi$, is defined as the angle between the normal to the geoid, normal to the limb ray path, and the equator. The radius of the orbit and the oblateness of the Earth have been exaggerated for clarity.

to acknowledge this interdependence, otherwise the retrieval process can exhibit unphysical behavior due to the double bookkeeping taking place. The way to get around this problem is to construct a set of fundamental coordinates for the system.

Clearly, the development so far indicates that pressure is the most appropriate fundamental coordinate in the vertical direction. The appropriate horizontal coordinate should account for the inclination of the EOS CHEM orbit and the oblateness of the Earth. It is desirable to use a coordinate system within which the MLS scans are relatively evenly spaced, as many of the quantities used in the forward model will be precomputed on regular horizontal and vertical grids. Clearly latitude is not an appropriate candidate, as the scan spacing is not constant in latitude. In fact, at high latitudes there is a degeneracy in the latitude coordinate, due to the inclination of the EOS CHEM orbit. For EOS MLS, the coordinate $\phi$, described in Figure 3.7, will be used as a fundamental coordinate.

# Chapter 4
# The EOS MLS Level 2 data processing algorithms

Section 3.6 discussed the details of the EOS MLS observation geometry. A key point is that radiance observations from one scan are dependent on the state of the atmosphere over a horizontal range of order of a few profile spacings. This fact provides a useful tool for characterizing the details of atmospheric variability along the spacecraft flight direction, if it is explicitly included in the calculations.

In order to take advantage of this, horizontal homogeneity cannot be assumed in either the forward model or retrieval processes. The approach taken is to retrieve the data in blocks of contiguous profiles. Such a calculation is not prohibitive, due to the comparative sparsity of the matrices involved. This Chapter considers in detail the implementation of the retrieval calculation in this manner.

Other aspects of the implementation of the retrieval calculation are also discussed, including improving the efficiency of the calculation, and its numerical stability.

## 4.1 A simple one dimensional approach

Before discussing the full two-dimensional MLS retrieval system described above, it is helpful to consider a simpler system. Equation 3.13, repeated here, gives an iterative expression for the retrieval operation

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} + \left[\mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1}\left\{\mathbf{S}_\mathbf{a}^{-1}\left[\mathbf{a} - \mathbf{x}^{(r)}\right] + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x}^{(r)})\right]\right\}.$$

This simple retrieval method is a one-dimensional approach, where all the issues of the observation geometry discussed in Section 3.6 are ignored. Instead the assumption is made that one complete scan of radiance observations depends only on one atmospheric profile, and that horizontal homogeneity can be assumed. This turns the retrieval calculation into $N$ separate individual profile retrievals, where $N$ is the number of profiles under consideration. This approach corresponds exactly to the UARS MLS version 5 retrieval algorithm.

As this simple algorithm ignores the geometrical issues raised in Section 3.6, it is not optimum for producing the final geophysical products. However it can be useful in obtaining a first guess state vector for use in a later 2D retrieval. It will transpire that this simple method plays an important role in the full retrieval algorithm.

## 4.2 Structure and sparsity in the full MLS retrieval system

In the full MLS retrieval system, profiles[1] will be retrieved in blocks of length $N$, where $N$ is expected to be of order 100, corresponding to a significant fraction of an orbit (1 orbit = 240 scans/profiles.) In order to overcome edge effects, the blocks will overlap slightly. In this scheme, the state vector will

[1] In this discussion the term 'profile' (singular) is taken to refer to a complete set of vertical profiles of temperature and composition for a *single* location on the globe.

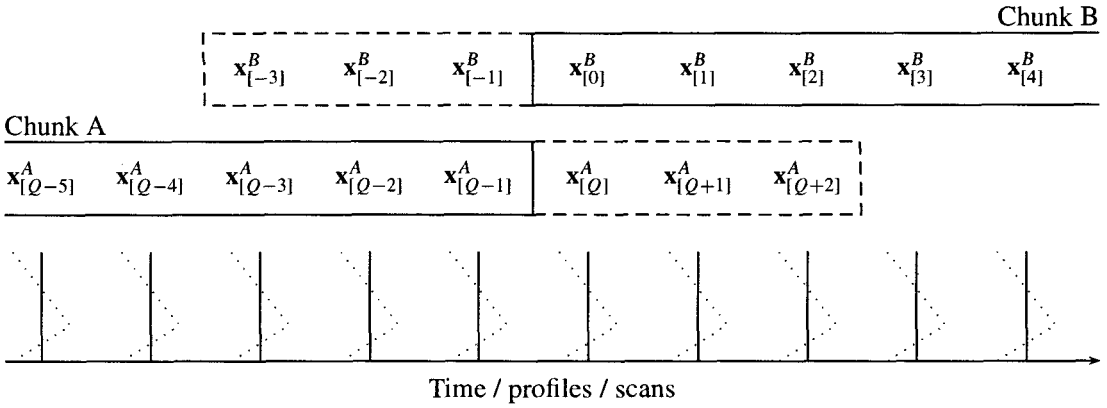



**Figure 4.1:** This figure shows an example of how the MLS profiles will be retrieved in chunks. Each chunk consists of $Q$ profiles, with an additional $q$ profiles of overlap (3 in this case). The total number of profiles in a chunk $N$ is thus $Q + 2q$. Beyond the end of each chunk, horizontal homogeneity is assumed, the overlaps are included (where possible) to reduce the impact of this approximation on the retrieved data. The data in the overlapping regions will be compared for diagnostic purposes. For example, $\mathbf{x}^A_{[Q]}$ should be similar to $\mathbf{x}^B_{[0]}$.

consist of $N$ profiles, with measurement vectors corresponding to $M$ scans. In most cases $N = M$, giving a one to one correspondence between scans and profiles. However, this is not a requirement, one could construct a state vector consisting of two profiles per scan, or one profile every two scans if necessary. Figure 4.1 shows the $N = M$ arrangement.

For the purposes of this discussion, consider a system with only one measurement vector. The results obtained here extend trivially into the multiple measurement vector case. The state vector consists of $N$ profiles $\mathbf{x}_{[j]}$ each of length $n$. There are additional state vector elements $\mathbf{x}^\diamond$ which contain any 'constant' quantities such as spectroscopy and instrument calibration (e.g. sideband ratios) which may be sufficiently uncertain that it is appropriate to retrieve them.

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}^\diamond \\ \mathbf{x}_{[0]} \\ \mathbf{x}_{[1]} \\ \vdots \\ \mathbf{x}_{[N-1]} \end{bmatrix} \tag{4.1}$$

The goal of the retrieval algorithm will then be to simultaneously obtain values for the entire vector $\mathbf{x}$. The main part of the retrieval calculation is the computation of $\left[\mathbf{S}_\mathbf{a}^{-1} + \mathbf{K}^\mathrm{T}\mathbf{S}_\mathbf{y}^{-1}\mathbf{K}\right]$ which is the topic of the next two subsections. The next section discusses the issues involved in inverting this matrix.

### 4.2.1 The weighting function matrices

Along similar lines to the construction of the state vector, the measurement vector consists of observations from $M$ scans. Section 3.6 showed that in the $N = M$ case, the measurements from one scan were affected by $\sim 5$ adjacent profiles, 2 either side of the nominal position. This amounts to saying that the $\mathbf{K}$ matrix has a bandwidth $p = 2$, such that $\mathbf{K}_{[\alpha j]} = \mathbf{0}$ for all $|\alpha - j| > p$, $j \neq 0$. As an illustration, consider a simpler case where $p = 1$, and $N = 6$, here the weighting function matrix has the block

structure

$$\mathbf{K} = \begin{bmatrix} \times & \times & \times & 0 & 0 & 0 & 0 \\ \times & \times & \times & \times & 0 & 0 & 0 \\ \times & 0 & \times & \times & \times & 0 & 0 \\ \times & 0 & 0 & \times & \times & \times & 0 \\ \times & 0 & 0 & 0 & \times & \times & \times \\ \times & 0 & 0 & 0 & 0 & \times & \times \end{bmatrix}. \tag{4.2}$$

The $\times$ symbol indicates a non-zero block sub matrix, while 0 indicates a sub matrix that is identically zero. Each row of the matrix corresponds to a separate scan. The first column indicates the weighting functions for the $\mathbf{x}^{\diamond}$ information, while the remaining columns indicate profiles $0 \cdots N-1$. This type of matrix is known as a singly-bordered (i.e. one column fully non zero) block band diagonal matrix.

### 4.2.2 The $\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\mathrm{y}}^{-1}\mathbf{K}$ matrix

As described in Sections 3.2.1 and 3.2.4 the measurement covariance matrix $\mathbf{S}_{\mathrm{y}}$ is diagonal. This gives a $\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\mathrm{y}}^{-1}\mathbf{K}$ matrix which is doubly-bordered block band diagonal with a block bandwidth of $2p$. The example $\mathbf{K}$ matrix of Equation 4.2 gives

$$\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\mathrm{y}}^{-1}\mathbf{K} = \begin{bmatrix} \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & 0 & 0 & 0 \\ \times & \times & \times & \times & \times & 0 & 0 \\ \times & \times & \times & \times & \times & \times & 0 \\ \times & 0 & \times & \times & \times & \times & \times \\ \times & 0 & 0 & \times & \times & \times & \times \\ \times & 0 & 0 & 0 & \times & \times & \times \end{bmatrix}. \tag{4.3}$$

### 4.2.3 The *a priori* covariance matrix

The *a priori* covariance matrix describes the uncertainty on the *a priori* estimate of the state vector, or a subset of the state vector. For the moment, consider a very simple state vector containing a set of $N$ vertical profiles of atmospheric temperature on $n$ surfaces. Define $T_{[(ij)]}$ to be the temperature for profile $i$ on surface $j$. The diagonal elements of the covariance matrix $\mathbf{S}_{\mathbf{a}}$ will describe the uncertainties on the individual temperatures. The off diagonal elements of $\mathbf{S}_{\mathbf{a}}$ describe the covariance between different profiles/surfaces according to

$$[\mathbf{S}_{\mathbf{a}}]_{(ij)(pq)} = \sqrt{[\mathbf{S}_{\mathbf{a}}]_{(ij)(ij)}\,[\mathbf{S}_{\mathbf{a}}]_{(pq)(pq)}}\; f(h_i, h_p, v_j, v_q), \tag{4.4}$$

where $h_i, h_p$ are the horizontal coordinates of profiles $i$ and $p$, and $v_j, v_q$ are the vertical coordinates of surfaces $j$ and $q$. The function $f$ describes correlation between adjacent profiles and surfaces. Clearly $f$ needs to be symmetric in the $h$ and $v$ terms, and should equal 1 if $h_i = h_p$ and $v_j = v_q$. Appropriate forms of $f$, such as exponential decay can guarantee that $\mathbf{S}$ is positive definite.

While it is a slight limitation on the flexibility of the system, extreme efficiency can be obtained if the forms of $f$ are restricted to those such that

$$f(h_i, h_p, v_j, v_q) = f_h(h_i, h_p)\, f_v(v_j, v_q), \tag{4.5}$$

i.e. cases where horizontal and vertical correlations are independent. Clearly, again, $f_h$ and $f_v$ have to be symmetric in their two arguments, and must evaluate to unity when their arguments are the same.

Along with this restriction, $S_{(ij)(ij)}$ is assumed to be independent of $i$, i.e. the temperature uncertainty is a function of height only, $[\mathbf{S_a}]_{(ij)(ij)} = \sigma_j^2$.

In this case $\mathbf{S_a}$ can be expressed as a block matrix according to

$$\mathbf{S_a} = \begin{bmatrix} H_{11}\mathbf{S_v} & H_{12}\mathbf{S_v} & \cdots & H_{1N}\mathbf{S_v} \\ H_{21}\mathbf{S_v} & H_{22}\mathbf{S_v} & \cdots & H_{2N}\mathbf{S_v} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N1}\mathbf{S_v} & H_{N2}\mathbf{S_v} & \cdots & H_{NN}\mathbf{S_v} \end{bmatrix}, \tag{4.6}$$

where the $N \times N$ matrix $\mathbf{H}$ is given by

$$\mathbf{H} = \begin{bmatrix} 1 & f_h(h_1, h_2) & \cdots & f_h(h_1, h_N) \\ f_h(h_2, h_1) & 1 & \cdots & f_h(h_2, h_N) \\ \vdots & \vdots & \ddots & \vdots \\ f_h(h_N, h_1) & f_h(h_N, h_2) & \cdots & 1 \end{bmatrix} \tag{4.7}$$

and the $n \times n$ matrix $\mathbf{S_v}$ is defined as

$$\mathbf{S_v} = \begin{bmatrix} \sigma_1^2 & f_v(v_1, v_2)\sqrt{\sigma_1^2\sigma_2^2} & \cdots & f_v(v_1, v_n)\sqrt{\sigma_1^2\sigma_n^2} \\ f_v(v_2, v_1)\sqrt{\sigma_2^2\sigma_1^2} & \sigma_2^2 & \cdots & f_v(v_2, v_n)\sqrt{\sigma_2^2\sigma_n^2} \\ \vdots & \vdots & \ddots & \vdots \\ f_v(v_n, v_1)\sqrt{\sigma_n^2\sigma_1^2} & f_v(v_n, v_2)\sqrt{\sigma_n^2\sigma_2^2} & \cdots & \sigma_n^2 \end{bmatrix}. \tag{4.8}$$

Comparison of Equation 4.8 with Equation 4.4 shows that $\mathbf{S_v}$ can be considered as the covariance matrix for an individual profile, with $\mathbf{H}$ as a matrix describing the horizontal correlation. The matrix construct in Equation 4.6 is called a *Kronecker product*

$$\mathbf{S_a} = \mathbf{H} \otimes \mathbf{S_v}. \tag{4.9}$$

The efficiency is gained by invoking the powerful identity

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}. \tag{4.10}$$

This allows for highly efficient computation and storage of $\mathbf{S_a}^{-1}$ in Equation 3.13. While the matrix is not strictly speaking sparse, the amount of information it contains is very small, given its size. This leads to a retrieval system that is comparatively simple to solve.

This result also holds for cases where the state vector is more complex, containing quantities of different types; provided one does not require there to be covariance between different families of *a priori* state vector elements, such as between temperature and ozone abundance.

### 4.2.4 Sparsity in the individual block sub matrices

As well as being sparse in the block sense, many of the individual submatrices involved in the retrieval calculation will themselves be sparse. However, only when the sparsity in a block submatrix is significant can appreciable savings be made. For example, the ∂[Radiance]/∂[Composition] submatrices will typically be about 50% zero, the possible savings to be made in not multiplying by zero in this case would easily be outweighed by the burden of storing and perusing the matrix in sparse form.

One set of block submatrices that will be highly sparse is those involving the tangent pressure $\vec{\zeta}$ quantity. Each radiance will only be dependent on the tangent pressure for its own MIF. Thus, for a 25

channel filter bank, with 120 MIFs, the full matrix size would be $(25 \times 120) \times 120 = 32\,320$, however, only $25 \times 120 = 3\,000$ (0.8%) of these values will be non zero. This is significant, as the $\vec{\zeta}$ quantity, having 120 elements, has far larger derivative matrices than say temperature where 40 – 50 elements per profile are anticipated.

## 4.3 Solving sparse matrix problems

While the matrices involved in Equation 3.13 have been shown to be sparse, this does not necessarily help matters, as the retrieval calculation requires the inverse of the $\left[\mathbf{S}_\mathbf{a}^{-1} + \mathbf{K}^\mathrm{T}\mathbf{S}_\mathbf{y}^{-1}\mathbf{K}\right]$ matrix to be computed. The inverse of a sparse matrix is not necessarily itself sparse. It is possible to adapt Cholesky decomposition (the standard method for inverting symmetric positive definite matrices) to optimize it for this sparse problem. However, as the inverse (and the decomposition) of the sum of two matrices is not the sum of the inverses (or decompositions); no advantage could be made of the Kronecker product structure in $\mathbf{S}_\mathbf{a}$. There are many efficiencies to be gained by invoking faster matrix solving methods based on iterative techniques.

### 4.3.1 The use of iterative matrix solvers

The aim of the retrieval problem is to compute a value for $\mathbf{x}^{(r+1)}$ according to Equation 3.13

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} + \left[\mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1} \left\{\mathbf{S}_\mathbf{a}^{-1}\left[\mathbf{a} - \mathbf{x}^{(r)}\right] + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x}^{(r)})\right]\right\}.$$

This expression can be simply rearranged to give

$$\left[\mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\mathbf{K}_i\right]\left[\mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}\right] = \mathbf{S}_\mathbf{a}^{-1}\left[\mathbf{a} - \mathbf{x}^{(r)}\right] + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x}^{(r)})\right], \tag{4.11}$$

which is equivalent to

$$\mathbf{M}\mathbf{z} = \mathbf{b}. \tag{4.12}$$

where

$$\mathbf{M} = \left[\mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\mathbf{K}_i\right], \tag{4.13}$$

$$\mathbf{z} = \mathbf{x}^{(r+1)} - \mathbf{x}^{(r)}, \text{ and} \tag{4.14}$$

$$\mathbf{b} = \mathbf{S}_\mathbf{a}^{-1}\left[\mathbf{a} - \mathbf{x}^{(r)}\right] + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x}^{(r)})\right], \tag{4.15}$$

An iterative matrix solving technique gives the value of $\mathbf{z}$ (sometimes called the *innovation* in $\mathbf{x}$) without needing to explicitly compute the value of $\mathbf{M}^{-1}$. This approach can be significantly more efficient than a full matrix inversion based technique.

The disadvantage of this method in the retrieval theory context is that, as Equation 3.14 showed, the $\mathbf{M}^{-1}$ matrix describes the covariance of the solution state vector. If this matrix is not calculated, no information on the estimated uncertainty of the solution is forthcoming.

This does not present too serious a problem. If the iterative matrix solver is sufficiently fast and accurate, the solution covariance matrix could be obtained by solving a series of $\mathbf{M}\mathbf{z} = \mathbf{b}$ problems where the $\mathbf{b}$ vectors were the successive columns of the identity matrix, while still being faster than a 'classical' matrix inversion method. There is a more appropriate solution to this problem which will be discussed shortly.

### 4.3.2 The conjugate gradient method

The conjugate gradient method (Golub and VanLoan 1996) is the most efficient and commonly used iterative matrix solving technique. The aim of the method is to solve the $\mathbf{Mz} = \mathbf{b}$ problem by seeking to minimize

$$\psi(\mathbf{z}) = \frac{1}{2}\mathbf{z}^{\mathrm{T}}\mathbf{Mz} - \mathbf{z}^{\mathrm{T}}\mathbf{b}. \tag{4.16}$$

The minimum value of $\psi = -\mathbf{b}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{b}/2$ is reached at $\mathbf{z} = \mathbf{M}^{-1}\mathbf{b}$. The name of the technique derives from the method by which the successive search directions are chosen.

In the simpler, steepest descent approach, separate one-dimensional minimizations are performed each iteration, along the line of steepest descent. This means that each search direction necessarily has to be orthogonal to the previous one, as the previous iteration has specifically chosen the location where the gradient along this previous search direction was zero. This can lead to very slow convergence for many cases, where the system is forced to traverse back and forth across a long narrow valley, rather than straight down. The problem is that the successive search directions are too different. The conjugate gradient method uses a more sophisticated method to choose search directions, based both on the previous search directions and the current residual ($\mathbf{Mz} - \mathbf{b}$) vector.

The reason that the conjugate gradient method is so suited to this particular problem, is that it never refers to matrix $\mathbf{M}$ directly; instead, each iteration a user-supplied function is invoked to compute the value of $\mathbf{Mz}$. This function can then be specially devised to take advantage of all the sparsity and structure in the $\mathbf{M}$ matrix discussed previously.

In this particular case, great efficiency can be gained by separating out the $\mathbf{Mz}$ calculation according to

$$\mathbf{Mz} = \left[\mathbf{S}_{\mathbf{a}}^{-1} + \sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{K}_i\right]\mathbf{z} = \mathbf{S}_{\mathbf{a}}^{-1}\mathbf{z} + \left[\sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{K}_i\right]\mathbf{z}. \tag{4.17}$$

The $\mathbf{S}_{\mathbf{a}}^{-1}\mathbf{z}$ calculation can take full advantage of the fact that $\mathbf{S}_{\mathbf{a}}^{-1}$ is a Kronecker product, while the $\left[\sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{K}_i\right]\mathbf{z}$ calculation takes advantage of the sparsity of $\left[\sum_i \mathbf{K}_i^{\mathrm{T}}\mathbf{S}_i^{-1}\mathbf{K}_i\right]$. Further efficiency and/or numerical stability may be achieved by breaking down the summation over $i$, though this will depend on the detailed sparsity of the matrices involved.

The method is guaranteed to produce the exact answer in $n$ iterations, where $n$ is the size of the $\mathbf{M}$ matrix. However, for some matrices, notably those with clustered eigenvalues, or 'close' to the identity matrix, convergence is much faster.

### 4.3.3 The use of preconditioning

The convergence properties described above lead to the concept of preconditioning. It is mathematically possible to rotate the retrieval system, such that the $\mathbf{M}$ matrix is 'closer' to the identity. As described above, this would lead to faster convergence of the conjugate gradient method. Of course, a full rotation computation would involve as much effort as solving the matrix completely, so little would be gained.

However, it transpires that rotating the system in this manner is mathematically equivalent to providing, each iteration, the solution to a different matrix equation $\mathbf{Pw} = \mathbf{c}$ where $\mathbf{P}$ is the *preconditioner* matrix. The preconditioner matrix is one that is 'similar' to $\mathbf{M}$. 'Similar' in this context is hard to define precisely, but a typical preconditioner matrix might be a sparser version of the matrix $\mathbf{M}$. The matrix $\mathbf{P}$ must be significantly simpler to solve than $\mathbf{M}$, otherwise the exercise is pointless.

In the case of the MLS retrieval system, an obvious candidate for the preconditioner system is the simple one dimensional retrieval approach described in Section 4.1. Other candidates include a

trimmed down version of the **M** matrix, with many off diagonal blocks removed, or a number of similar retrievals to this full one with smaller values of $N$. It is also possible to nest retrieval algorithms with the preconditioner itself being preconditioned by a third retrieval algorithm.

The preconditioner can also be used to obtain the solution covariance matrix, as being the inverse of **P**, rather than of **M**. As it is typically only diagonal elements of this covariance matrix that are reported, the approximation involved is comparatively small.

The full derivation of the preconditioned conjugate gradient calculation is well established (see Golub and VanLoan 1996 or other texts,) for completeness, the details are shown in Algorithm 4.1. This algorithm has been applied to many different calculations. In particular, the Goddard Space Flight Center Data Assimilation Office (DAO), use this method to solve the matrix equation that forms the basis of their assimilation calculation (Data Assimilation Office 1996).

---

**Algorithm 4.1:** The preconditioned conjugate gradient method. Given a symmetric positive definite $n \times n$ matrix **M**, $n$ element vectors **b** and initial guess $\mathbf{z}_{(0)}$, the following algorithm solves the linear system $\mathbf{Mz} = \mathbf{b}$, invoking another matrix solver for the preconditioner matrix **P**. The parenthetical subscripts indicate values from different iterations of the algorithm.

$k = 0$
$\mathbf{r}_{(0)} = \mathbf{b} - \mathbf{A}\mathbf{z}_{(0)}$
***while*** $\left[\mathbf{r}_{(k)} \neq \mathbf{0}\right]$
    *Solve* $\mathbf{P}\mathbf{w}_{(k)} = \mathbf{r}_{(k)}$ for $\mathbf{w}_{(k)}$.
    $k = k + 1$
    ***if*** $k = 1$
        $\mathbf{p}_1 = \mathbf{w}_{(0)}$
    ***else***
        $\beta_{(k)} = \mathbf{r}^{\mathrm{T}}_{(k-1)}\mathbf{w}_{(k-1)} / \mathbf{r}^{\mathrm{T}}_{(k-2)}\mathbf{w}_{(k-2)}$
        $\mathbf{p}_{(k)} = \mathbf{w}_{(k-1)} + \beta_{(k)}\mathbf{p}_{(k-1)}$
    ***end***
    *Compute* $\mathbf{q}_{(k)} = \mathbf{M}\mathbf{p}_{(k)}$
    $\alpha_{(k)} = \mathbf{r}^{\mathrm{T}}_{(k-1)}\mathbf{w}_{(k-1)} / \mathbf{p}^{\mathrm{T}}_{(k)}\mathbf{q}_{(k)}$
    $\mathbf{z}_{(k)} = \mathbf{z}_{(k-1)} + \alpha_{(k)}\mathbf{p}_{(k)}$
    $\mathbf{r}_{(k)} = \mathbf{r}_{(k-1)} - \alpha_{(k)}\mathbf{q}_{(k)}$
***end***
$\mathbf{z} = \mathbf{z}_{(k)}$

---

## 4.4 Increasing efficiency in the retrieval calculation

Even when taking maximum advantage of the sparsity of the matrices involved, the MLS retrieval algorithm can be a time-consuming calculation. The aim of this section is to consider various approaches to improving the speed of the calculation, and to outline the manner in which decisions would be arrived at regarding which compromises are appropriate to achieve a given saving.

### 4.4.1 Operation counts

The key to the issue of improving the speed of the retrieval calculations is the number of individual scalar operations required by the algorithm. However, the number of operations required by a calculation does

**Table 4.1:** A summary of the operation counts required for each outer iteration of the retrieval method. The $\nu$ symbol indicates the number of iterations of the preconditioned conjugate gradient method required. The 'independent' method retrieves the $N$ profiles independently. The 'full' and 'sparse' methods both retrieve the profiles in blocks of $N$, the sparse method taking advantage of the simplicity of the system. The $\mathbf{S_y}$ matrix is assumed diagonal, if this is not the case, the first operation increases by a factor of at least $m$ (possibly $Nm$.)

| Operation | Independent | Full | Sparse |
|---|---|---|---|
| $\mathbf{K}^T\mathbf{S_y}^{-1}$ | $\sum_i Nnm_i$ | $\sum_i N^2nm_i$ | $\sum_i Npnm_i$ |
| $\mathbf{K}^T\mathbf{S_y}^{-1}\mathbf{K}$ | $\sum_i Nn^2m_i$ | $\sum_i N^3n^2m_i$ | $\sum_i Np^2n^2m_i$ |
| solve $\left[\mathbf{S_a}^{-1} + \mathbf{K}^T\mathbf{S_y}^{-1}\mathbf{K}\right]$ | $Nn^3$ | $N^3n^3$ | $\nu Npn^2$ |
| $\mathbf{S_a}^{-1}[\mathbf{a}-\mathbf{x}] + \mathbf{K}^T\mathbf{S_y}^{-1}[\mathbf{y}-\mathbf{f}]$ | $Nn^2 + Nnm$ | $N^2n^2 + N^2nm$ | $Npn^2 + Npnm$ |

not directly relate to the time it takes to perform the calculation on a computer. Issues such as cache filling, parallel processing etc. have a big impact on the efficiency of a calculation. The operation counts serve merely as a 'rule of thumb' measure of the size of the task involved.

Table 4.1 gives a rough summary the number of operations required for the various stages in the EOS MLS retrieval algorithm. The summary assumes no sparsity within the profile / scan blocks at either the submatrix level (e.g. no allowance for the fact that $\partial$[R1A:118.B1F:PT]/$\partial$[ClO] = 0) or within the submatrices (e.g. the issue with $\vec{\zeta}$ discussed in Section 4.2.4.)

In the MLS case $m_i \gg n$ for all the measurement vectors. From this it is clear that the ability to assume that $\mathbf{S}_i$ is diagonal is crucial if the algorithm is to be efficient. This factor rules out the use of constrained quantity error propagation as explained in Section 3.2.4. It is clearly more efficient to attempt to retrieve the whole state vector from every measurement vector simultaneously, as opposed to proceeding in separate phases and propagating errors for the constrained quantities.

Given that the $\mathbf{S}_i$ matrix will be diagonal, the most computationally intensive step will be the construction of the $\mathbf{K}_i^T\mathbf{S}_i^{-1}\mathbf{K}_i$ matrices, taking $Np^2n^2m_i$ operations for each measurement vector. There are several points of interest to note about this calculation. Firstly, it is linear in $N$, that is to say, (ignoring the overlaps) retrieving 200 profiles in one single chunk will involve the same computational effort as retrieving two 100 profile chunks. This means that the size of $N$ is limited only by the memory capacity of the computer used. Secondly, while the operation is linear in $m_i$, it is quadratic in $n$ and $p$. While $p$, being geometrical is hard to control, $n$ can be changed easily, being a factor of the vertical resolution and range of the state vector quantities. Being a quadratic term, a factor of two increase in performance can result from removing only 30% of the elements of $\mathbf{x}$. $m_i$ can also be changed fairly easily by removing or combining radiances that contribute little information to the system.

### 4.4.2 Phasing revisited

Although the use of constrained quantity error propagation has been ruled out on the grounds of being too computationally intensive, this does not necessarily rule out the use of retrieval phasing. Phasing can still be a very useful tool in the algorithm. However, unlike in previous descriptions, the previously retrieved quantities are not constrained, rather new items are added to the pool of retrieved quantities.

In this manner, the more non-linear quantities (tangent pressure and temperature in particular), that will require more iterations for convergence can be retrieved alone in the earlier phases. The later phases

include quantities such as composition which will require a smaller number of iterations (1 or 2.) The non-linear quantities are still retrieved in these later phases; however, as convergence was achieved earlier, only small revisions will be made. By using this scheme, the number of iterations required for the full system is minimized.

Taking this one stage further, it is also possible in the earlier phases of the retrieval process to constrain quantities without propagating errors for them, or assuming that the modified $\mathbf{S}_i$ matrices are diagonal. As these early stages are only used to find a suitable starting point for the final 'full' retrieval, there is no need for a comprehensive treatment of the error budget.

It is even possible to conceive of more complex systems whereby the temperature is retrieved at full resolution in an early phase, and the results output as the standard temperature product. In the later phases, where composition is the main goal, and temperature is less important (and has little impact on the radiances), the vertical resolution and/or range of the temperature profile could be decreased in the state vector, in order to reduce the computation effort.

### 4.4.3 An 'Information' perspective on retrieval issues

The previous sections touched on the mechanisms whereby the retrieval algorithm can be made more efficient, namely, reducing the size of $n$ or $m_i$. Clearly, some metric is needed to indicate which reduction schemes are preferable. One such metric is the *information content* of the retrieval system.

The information content of a system is a measure of the size (in fact 'inverse size') of the region in state space within which one is confident that the system is located. This region is described by the covariance matrix, which describes a hyperellipsoidal region in state space. The volume of this region is given by the product of the lengths of all the axes of the ellipsoid (give or take factors of $\pi$). These lengths are given by the square roots of the eigenvalues of the covariance matrix (recall that a covariance matrix is an inherently 'squared' quantity.) Thus the volume of state space is given by the square root of the product of the eigenvalues of the covariance matrix. Recall that the determinant of a matrix ($|\cdots|$) is equal to this product.

The information content ($H$) of a system is thus defined by the logarithm of the reciprocal of this volume, i.e.

$$H = \log_2 \left[\sqrt{|\mathbf{S}|}\right]^{-1} = -\frac{1}{2}\log_2 |\mathbf{S}| = \frac{1}{2}\log_2 \left|\mathbf{S}^{-1}\right|. \tag{4.18}$$

(Recall that $|\mathbf{M}|^{-1} = \left|\mathbf{M}^{-1}\right|$.) Base 2 is typically used for the logarithm, so that the information content can be described in terms of the number of 'bits' of information available.

In studies of retrieval systems it is useful to consider the information content added by the retrieval operation. Combining Equation 4.18 with Equation 3.14 describing the covariance of the retrieved product gives the following expression for the information added by a retrieval calculation (recall that $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$):

$$\Delta H = [\text{Final information content}] - [a\ priori \text{ information content}] \tag{4.19}$$

$$= \frac{1}{2}\log_2 \left|\mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^T \mathbf{S}_i^{-1} \mathbf{K}_i\right| - \frac{1}{2}\log_2 \left|\mathbf{S}_\mathbf{a}^{-1}\right| \tag{4.20}$$

$$= \frac{1}{2}\log_2 \left|\mathbf{I} + \left[\sum_i \mathbf{K}_i^T \mathbf{S}_i^{-1} \mathbf{K}_i\right] \mathbf{S}_\mathbf{a}\right|. \tag{4.21}$$

In practice, the determinant operation is highly numerically unstable, so it is best to apply a sequential approach to the calculation of the information content, adding measurements one at a time. The derivation of this calculation is somewhat complex; see Appendix C.2 for more details.

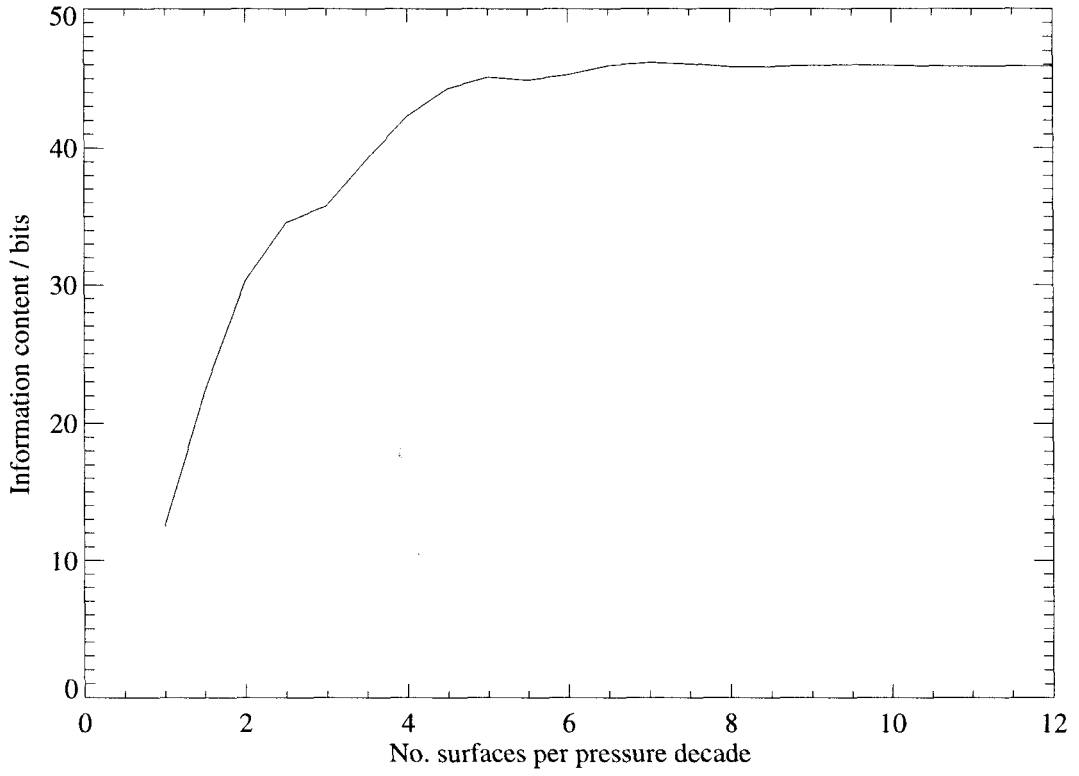



**Figure 4.2:** This plot shows the trade-off between resolution and information content for a system measuring $O_3$ from the `R2:190.B6F:O3` radiance observations (temperature and pressure are taken to be perfectly known.) The $O_3$ profile spans 1000 to 0.01 mb.

### 4.4.4 Vertical resolution

Given that the rate determining step in the retrieval calculation is order $n^2m$, cutting down the size of the state vector will give a significant increase in speed. The main way in which to decrease this size is to decrease the vertical resolution of the geophysical profiles in the state vector. However, too coarse a resolution will give a deterioration in information content. Figure 4.2 shows the trade-off between the vertical resolution and information gained for the ozone observations from band `R2:190.B6F:O3`.

The figure shows that little information is lost by going from twelve to six surfaces per decade for the $O_3$ profile. Given that this is a factor of two decrease in the state vector length, a quadrupling of speed could be seen if this were implemented. If similar savings can be made in all the other species, significant speed increases could be obtained. In addition, the vertical range of the profiles could be limited, or the resolution further degraded in certain regions of the profile (for example where the signal to noise ratio is poor.)

### 4.4.5 Radiance averaging / limiting

In addition to decreasing the resolution and / or range of the state vector components, the number of radiances used in the retrieval calculations can be reduced by either limiting the vertical range of the radiances used, or combining radiances from adjacent minor frames. Again, the information content of the retrieval system is the metric whereby decisions would be made regarding the most appropriate

strategy. The form of the averaging or limiting would clearly vary from channel to channel, as the tangent point altitude range over which useful signals are obtained varies from channel to channel (see, for example, Figure 3.2.)

As a test case, a simple 'random walk' type algorithm has been implemented to determine the most advantageous scheme for a given reduction in the number of radiances used. Figure 4.3 shows the application of this algorithm to the `R2:190.B6F:O3` $O_3$ observations for the case where a 30% reduction in the effective number of radiances is required.

This example shows that rather than merging radiances from adjacent minor frames, the best way to reduce the number of radiances used in the retrieval calculation is to use the radiances at the full vertical resolution available, but over a limited vertical range which varies from channel to channel.

There are other points of interest to note from this example. For example, while the channel closest to the line center would be thought to give the most information about the upper regions of the atmosphere, the calculation has chosen to concentrate on the information from the two pairs of channels further away from the line center. This is due to the fact that the line center channel has a smaller band width and therefore a poorer signal to noise ratio than the channels further out. Figure 4.4 shows the optimum information content achieved by this search as a function of the fraction of the number of radiances used in the retrieval calculation.

In addition to averaging together radiances from multiple minor frames, it is also possible to average together radiances from different channels. While this is not discussed in detail here, a similar approach to the one outlined above can be used to arrive at optimum averaging schemes. This technique will be particularly useful when using radiance observations from the digital autocorrelator spectrometers (DACs). These spectrometers provide ~2 MHz resolution over ~10 MHz, for each minor frame. The amount of information supplied by these measurements is very small compared to the data volume. Techniques such as these allow the autocorrelator data to be effectively used.

Note that reducing the size of the state vector can also increase the efficiency of the forward model calculation. However, reducing the number of radiances used does not necessarily speed up the forward model, as the field of view convolution calculation in the forward model requires radiances at relatively high resolution over a large vertical range, independent of how many radiances are actually required for the retrieval calculation.

### 4.4.6 Implementation of these schemes in the production processing

Decisions made on the resolution of the state vector, and the reduction in size of the measurement vectors, will typically be made with reference to more specifically targeted quantities than those described here. For example, one might choose to maximize the information obtained in the lower stratosphere/upper troposphere region, at the expense of less precision in the mesosphere and upper stratosphere.

Additionally, note that these simple studies considered the retrieval of one molecule from a single band. Decisions made about the configuration of the production retrieval algorithms will be made with reference to studies of the complete measurement system.

## 4.5 Numerical stability considerations

The issue of numerical stability is very important in these calculations. The contents of the state vector represents a huge dynamic range, consider the contrast between the dynamic ranges of temperature (~40 K) and ClO mixing ratio (~ $1 \times 10^{-9}$). Care must be taken to avoid numerical round off errors when combining such quantities.

**(a)** Initial down-sample scheme

**(b)** Final down-sample scheme

**Figure 4.3:** (color) This figure shows the outcome of a search for the 'optimum' radiance down-sampling scheme for a desired 30% reduction in the number of radiances used in a retrieval of ozone at six surfaces per decade from `R2:190.B6F:O3`. These plots show different schemes for averaging together radiance measurements in order to reduce the size of the measurement vector. Where radiances have been averaged together, they are connected with a line. Those radiances that have not been averaged are represented by a point.
The left hand figure shows the initial scheme used, with radiances being averaged together in groups of about three to achieve the 30% reduction in measurement vector size. The right hand figure shows the scheme chosen based on maximizing the information content of the resulting system, while keeping this 30% cut.
The maximum possible information content, were all the radiances used is 45.29 bits, the information content for the initial system (left hand figure) was 34.59 bits. The final information content achieved for the optimal scheme was 38.22 bits.
Only the first 13 channels have been shown, the radiances curves have been successively offset by 10 K for clarity.

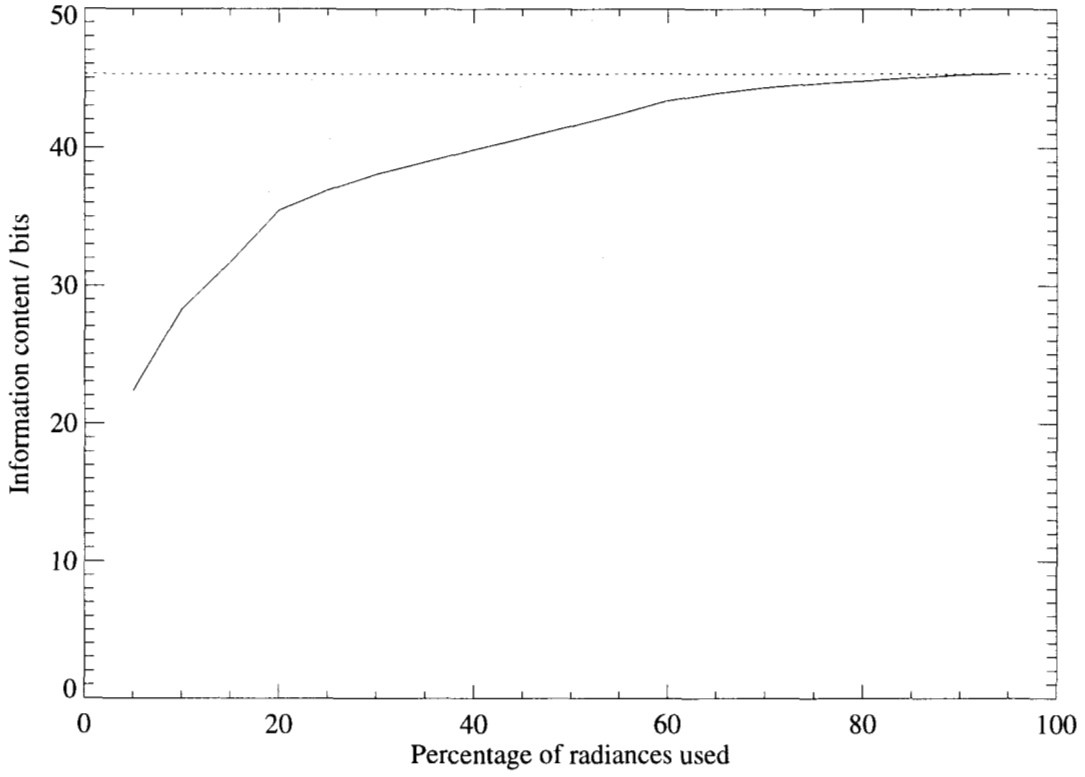



**Figure 4.4:** This figure shows the trade-off between the number of radiances used in the `R2:190.B6F:O3` ozone retrieval (at six surfaces per decade) and the optimum information content that can be obtained from the retrieval system using the given number of radiances.

This problem can be avoided by scaling the system, to normalize all the quantities involved. This normalization process yields 'scaled' vectors according to

$$\widetilde{\mathbf{x}} = \mathbf{S}_{\mathbf{a}}^{-\frac{1}{2}}\mathbf{x} \tag{4.22}$$

$$\widetilde{\mathbf{y}}_i = \mathbf{S}_i^{-\frac{1}{2}}\mathbf{y}_i, \tag{4.23}$$

where the superscript $-\frac{1}{2}$ indicates the inverse square root operation for the matrix (i.e. the inverse of its Cholesky Decomposition, see Golub and VanLoan 1996; Rodgers 1996). The $\mathbf{K}_i$ matrices are scaled according to

$$\widetilde{\mathbf{K}}_i = \mathbf{S}_i^{-\frac{1}{2}}\mathbf{K}_i\mathbf{S}_{\mathbf{a}}^{\frac{1}{2}}. \tag{4.24}$$

The retrieval equations become simpler in this scaled space because the covariance matrices $\widetilde{\mathbf{S}}_{\mathbf{a}}$, and $\widetilde{\mathbf{S}}_i$ are all identity matrices.

As has been stated previously, the measurement covariance matrices $\mathbf{S}_i$ will typically be diagonal. The *a priori* covariance matrix $\mathbf{S}_{\mathbf{a}}$ however, presents more of a challenge. As described in Section 4.2.3, the matrix is a Kronecker product, so computing its inverse (or in this case its Cholesky decomposition) is trivial. However, applying the scaling of $\mathbf{K}_i$ from Equation 4.24 gives a $\widetilde{\mathbf{K}}_i$ with less sparsity than the original matrix, due to the inter-profile correlation imposed by the *a priori* covariance. This would significantly increase the required computation time.

A more efficient approach is to use a simpler matrix than $\mathbf{S_a}$ as the basis for the state vector scaling. Consider

$$\widetilde{\mathbf{x}} = \mathbf{S_p}^{-\frac{1}{2}}\mathbf{x}, \tag{4.25}$$

where

$$\mathbf{S_p} = \mathbf{I}_N \otimes \mathbf{S_v}, \tag{4.26}$$

where $\mathbf{S_v}$ is as defined in Section 4.2.3, and $\mathbf{I}_N$ is the $N \times N$ identity matrix. This is essentially the same as $\mathbf{S_a}$ with all the off-diagonal blocks set to zero, and thus describes a system with no horizontal correlations.

The advantage of using such a matrix is that it performs the normalization required for stability while both still being trivial to decompose, and maintaining the sparsity in $\widetilde{\mathbf{K}}_i$. Applying this scaling gives $\widetilde{\mathbf{S}}_\mathbf{a} = \mathbf{H} \otimes \mathbf{I}_n$. All the major equations derived up to this point are presented in their scaled form in Table 4.2.

## 4.6 Testing for suitable convergence

The proposed algorithm is an iterative non-linear minimization of a cost function ($\chi^2$) in which each iteration itself involves the use of an iterative method for solving the matrix system. Clearly, both of these sets of iterations require a convergence test of some form.

As in most cases, these tests will be implemented by examining the size of the change in the state vector $\mathbf{x}$ each iteration. If the state vector has not changed significantly, according to some user defined threshold, then the iterations are deemed to have converged.

As discussed in the previous section, the MLS state vector contains many diverse quantities representing very different orders of magnitude. For this reason, it is preferable to perform the convergence test on the scaled state vector $\widetilde{\mathbf{x}}$, rather than directly on the state vector $\mathbf{x}$. This will weight the changes in each state vector element according to the *a priori* knowledge of their variability.

The convergence criteria will typically be a threshold for some norm of the change in $\widetilde{\mathbf{x}}$, for example the $L_2$ norm (i.e. $|\Delta\widetilde{\mathbf{x}}|$), or the $L_\infty$ norm (the maximum value).

## 4.7 Summary of proposed algorithm

For completeness the algorithm described above is summarized here. For simplicity only the case of Newtonian iteration is shown, as the Marquardt Levenberg method's requirements for additional decisions and a memory of previous states complicates the system. It is intended that the Marquardt-Levenberg method will be available in the MLS Level 2 data processing algorithms.

The paragraphs below summarize the actions taken during such an iterative retrieval, given an initial value for the state vector $\mathbf{x}$, a set of measurement vectors $\mathbf{y}_i$ with associated covariance matrices $\mathbf{S}_i$, and an *a priori* state vector $\mathbf{a}$ with its covariance matrix given in terms of $\mathbf{S_v}$ and $\mathbf{H}$ as described in Section 4.2.3.

1. The first stage of the retrieval process is (assuming it has not already been performed) the scaling of the measurement vectors $\mathbf{y}_i$ by the inverse of their covariance matrices. In other words

$$\widetilde{\mathbf{y}}_i = \mathbf{S}_i^{-\frac{1}{2}}\mathbf{y}_i \tag{4.30}$$

| Equation | Unscaled form | Scaled form | |
|---|---|---|---|
| 3.10 | $\mathbf{K}_i = \frac{\partial \mathbf{f}_i(\mathbf{x})}{\partial \mathbf{x}}$ | $\widetilde{\mathbf{K}}_i = \frac{\partial \widetilde{\mathbf{f}}_i(\widetilde{\mathbf{x}})}{\partial \widetilde{\mathbf{x}}}$ | (4.27) |
| 3.13 | $\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} + \left[\mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1} \left\{\mathbf{S}_\mathbf{a}^{-1}\left[\mathbf{a} - \mathbf{x}^{(r)}\right] + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x}^{(r)})\right]\right\}$ | $\widetilde{\mathbf{x}}^{(r+1)} = \widetilde{\mathbf{x}}^{(r)} + \left[\mathbf{H}^{-1} \otimes \mathbf{I}_n + \sum_i \widetilde{\mathbf{K}}_i^\mathrm{T}\widetilde{\mathbf{K}}_i\right]^{-1} \left\{\left[\mathbf{H}^{-1} \otimes \mathbf{I}_n\right]\left[\widetilde{\mathbf{a}} - \widetilde{\mathbf{x}}^{(r)}\right] + \sum_i \widetilde{\mathbf{K}}_i^\mathrm{T}\left[\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{f}}_i(\mathbf{x}^{(r)})\right]\right\}$ | (4.28) |
| 3.14 | $\mathbf{S}_\mathbf{x} = \left[\mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1}$ | $\widetilde{\mathbf{S}}_\mathbf{x} = \left[\mathbf{H}^{-1} \otimes \mathbf{I}_n + \sum_i \widetilde{\mathbf{K}}_i^\mathrm{T}\widetilde{\mathbf{K}}_i\right]^{-1}$ | (4.29) |

**Table 4.2:** This table shows scaled equivalents of various equations in this document. The scalings applied are those given in Equations 4.25 and 4.23, namely $\widetilde{\mathbf{y}}_i = \mathbf{S}_i^{-\frac{1}{2}}\mathbf{y}_i$, and $\widetilde{\mathbf{x}} = \mathbf{S}_\mathbf{p}^{-\frac{1}{2}}\mathbf{x}$, where $\mathbf{S}_\mathbf{p} = \mathbf{I}_N \otimes \mathbf{S}_\mathbf{v}$.

In nearly all cases $\mathbf{S}_i$ will be a diagonal matrix, so no difficulty is expected with this step. However, the algorithm will have the capability to deal with non-diagonal $\mathbf{S}_i$ matrices should the need arise for research purposes. Typically this will take the form of a standard Cholesky Decomposition algorithm.

In addition, the $\mathbf{x}$ and $\mathbf{a}$ vectors will be scaled according to

$$\widetilde{\mathbf{a}} = [\mathbf{I}_N \otimes \mathbf{S}_\mathrm{v}]^{-\frac{1}{2}}\,\mathbf{a}, \text{ and } \widetilde{\mathbf{x}} = [\mathbf{I}_N \otimes \mathbf{S}_\mathrm{v}]^{-\frac{1}{2}}\,\mathbf{x} \tag{4.31}$$

2. Given the current state vector, the forward model is then invoked to compute predicted radiances $\mathbf{f}_i$ corresponding to the observations $\mathbf{y}_i$, and associated weighting function matrices $\mathbf{K}_i$. It is possible that the weighting functions may not be recomputed, and simply assumed constant from one iteration to the next, in order to save computation time.

   In addition, if a preconditioner is being used, the forward model may also provide weighting function matrices corresponding to the preconditioner system (e.g. weighting functions corresponding to a 1D retrieval.)

3. Next the forward model radiance vectors are scaled, similarly to the measurement vectors $\mathbf{y}_i$ above.

$$\widetilde{\mathbf{f}}_i = \mathbf{S}_i^{-\frac{1}{2}}\mathbf{f}_i \tag{4.32}$$

   Additionally, the weighting function matrices $\mathbf{K}_i$ will be scaled (if they have been changed since the previous iteration), according to.

$$\widetilde{\mathbf{K}}_i = \mathbf{S}_i^{-\frac{1}{2}}\mathbf{K}_i\,[\mathbf{I}_N \otimes \mathbf{S}_\mathrm{v}]^{\frac{1}{2}} \tag{4.33}$$

   Once again, the fact that $\mathbf{S}_i$ is diagonal simplifies this process, as does the Kronecker Product construct, if it is properly exploited in the calculation.

   Any preconditioner system being used will also be scaled in a similar manner.

4. At this point the $\mathbf{M}$ matrices and $\mathbf{b}$ vectors are computed according to

$$\mathbf{M} = \left[\mathbf{H}^{-1} \otimes \mathbf{I}_n + \sum_i \widetilde{\mathbf{K}}_i^\mathrm{T}\widetilde{\mathbf{K}}_i\right] \tag{4.34}$$

$$\mathbf{b} = \left[\mathbf{H}^{-1} \otimes \mathbf{I}_n\right]\left[\widetilde{\mathbf{a}} - \widetilde{\mathbf{x}}^{(r)}\right] + \sum_i \widetilde{\mathbf{K}}_i^\mathrm{T}\left[\widetilde{\mathbf{y}}_i - \widetilde{\mathbf{f}}_i(\mathbf{x}^{(r)})\right] \tag{4.35}$$

   If any $\mathbf{K}_i$ matrices are unchanged from the previous iteration their contributions to $\mathbf{M}$ need not be recomputed (note that this implies a memory of the results of the various $\widetilde{\mathbf{K}}_i^\mathrm{T}\widetilde{\mathbf{K}}_i$ matrices). Additionally, once again, the $\mathbf{M}$ matrix will also be computed for any preconditioner systems.

5. The program will then solve $\mathbf{Mz} = \mathbf{b}$ for $\mathbf{z}$ which is equal to the innovation in $\widetilde{\mathbf{x}}$. Typically this will be by the use of a preconditioned conjugate gradient technique. In the case of the preconditioner, the $\mathbf{M}$ matrix will be explicitly decomposed and/or inverted as a preliminary step, in order to provide the solutions to the 'main' system upon demand.

6. The scaled state vector will be modified according to

$$\widetilde{\mathbf{x}} \rightarrow \widetilde{\mathbf{x}} + \mathbf{z}. \tag{4.36}$$

The result will then be unscaled according to

$$\mathbf{x} = [\mathbf{I}_N \otimes \mathbf{S}_\mathbf{v}]^{\frac{1}{2}} \widetilde{\mathbf{x}}. \tag{4.37}$$

If convergence test indicates insufficient convergence, the program will iterate, by returning to stage 2.

7. Finally, the unscaled state vector is output, and the final solution covariance is obtained as the inverse of **M** (possibly computed directly or by preconditioned conjugate gradients) or the inverse of the **M** matrix used in the preconditioner.

# Chapter 5
# Related algorithms for EOS MLS 'noisy' products

## 5.1 Introduction

Some of the molecules EOS MLS is designed to observe have particularly small mixing ratios, and weak emission lines. The corresponding radiance observations will thus have poor signal to noise ratios, leading to noisy retrievals. For these products, more useful results can be obtained by considering averaged products, such as daily zonal means, or monthly maps. There are several ways in which to compute such quantities.

## 5.2 Possible approaches

One approach is to simply retrieve the products in the same manner as all the others, and then use whatever averaging technique is appropriate afterwards. The disadvantage of this method is that, unless special care is taken, the *a priori* information can significantly bias the results, as it is included in each separate retrieval.

A second approach is to average the radiances from the relevant bands in whatever manner is appropriate, and to then perform retrievals on the averaged radiances. This method has a profound problem however when the lines of interest are contaminated by strong emission from other, highly variable molecules. This is the case for example with the BrO observations in `R4:640.B31M:BrO`, which are very close to a strong $O_3$ line.

From a computational point of view, the first approach does not represent any efficiency gain. The second method however can result in improved performance, as the radiances specifically targeted to the molecules in question can be ignored in the processing for each day.

## 5.3 The approach chosen

The best approach to this problem is to retrieve the averaged products as a separate task, after the main processing has occured. Rather than using averaged radiances as above, however, the full radiance data set for the relevant band is considered. Consider the iterative retrieval expression given in Equation 3.13

$$\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} + \left[\mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1} \left\{\mathbf{S}_\mathbf{a}^{-1}\left[\mathbf{a} - \mathbf{x}^{(r)}\right] + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\left[\mathbf{y}_i - \mathbf{f}_i(\mathbf{x}^{(r)})\right]\right\}.$$

In the linear (i.e. single iteration, with initial guess $\mathbf{x} = \mathbf{a}$) case, this reduces to

$$\mathbf{x} = \left[\mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1} \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\left[\mathbf{y}_i - \mathbf{f}_i\left(\mathbf{a}\right)\right]. \tag{5.1}$$

Now for the case of the noisy products take $\mathbf{x}$ to be a specific component of an averaged dataset (e.g. a single profile corresponding to one latitude in a daily zonal mean retrieval). Consider the measurement vectors $\mathbf{y}_i$ to represent each individual scan in the relevant spectral band that contributes to this component (e.g. all the scans in the latitude range under consideration.) The forward models for each scan use the previously retrieved values for the other molecules and parameters that affect the radiance measurements ($O_3$, temperature, tangent pressure etc.) as constrained quantities.

Note that the measurement covariance matrices used here are not necessarily diagonal. Firstly, as discussed in Section 3.5.4, the radiance observations may contain correlated errors due to the effects of $1/f$ noise. Furthermore, the effect of possible errors on the constrained quantities (temperature, tangent pressure etc.) need to be propagated through to the $\mathbf{S}_i$ matrices for these calculations. However, as the measurement vectors will only represent individual major frames from single bands (as opposed to chunks of several major frames for multiple bands), such a calculation is not prohibitive.

In this case $n$ (the size of $\mathbf{x}$) will be sufficiently small that traditional, non sparse, matrix computation methods will be appropriate. This system will probably be sufficiently linear that only one iteration is required.

# Chapter 6
# Algorithms for cloud flags and products

One of the advantages of the MLS instrument, when compared to infrared, visible, and ultraviolet instruments, is the ability to make atmospheric observations in the presence of clouds such as cirrus and polar stratospheric clouds (PSCs) and in the presence of dense aerosols formed from volcanic injections of $SO_2$. Because the observing wavelengths are significantly larger than the typical size of the cloud particles, clouds play a less important role in microwave remote sounding techniques than in infrared, visible and ultraviolet techniques. However, thick dense clouds in the troposphere can affect the MLS radiances and the effects of cloud on MLS observations in the upper troposphere need careful consideration.

## 6.1 The effect of clouds on MLS radiance observations

Figure 6.1 shows a radiance profile observed by UARS MLS at 204 GHz, bounded by two calculated profiles under clear-sky dry and supersaturation (110% relative humidity) conditions. At high tangent altitudes (above 8 km or ~300 hPa), clear-sky radiances are not expected to exceed the values defined by the supersaturated profile. However, excessive radiances can be induced by clouds through scattering and emitting upwelling radiation. One can use the supersaturated profile as a threshold to classify clear- and cloudy-sky observations at high tangent heights.

At lower tangent heights (below ~8 km or ~300 hPa in UARS case), there is a minimum defined by both dry and saturated profiles, which shows that the clear-sky radiances should not be less than these values. However, clouds can reduce the limb radiance significantly below the clear-sky limit by scattering some of upwelling radiation out of MLS field-of-view. This occurs when large (greater than $\sim 100\,\mu$m) ice particles are present, which can be the case with deep convective clouds. Thus, one can use the minimum radiances defined by the dry and supersaturated profiles to flag the low-altitude radiances that are affected by clouds.

Figure 6.2 shows a set of radiances observed by the UARS MLS 204 GHz channel. The smooth lines in the figure illustrate the approximate upper and lower limits defined by clear-sky dry and saturated conditions. Radiances outside this range have been affected by the presence of clouds. For each measured radiance one can calculate the corresponding clear-sky limits from an estimated temperature profile and use them for cloud detection thresholds.

The UARS MLS 204 GHz observations can be used to illustrate how such a method will be implemented for EOS MLS retrievals. Figure 6.3 shows the radiance difference between observed and maximum clear-sky radiances. Above ~300 hPa, most radiances are close to or less than zero with a few exceptions where radiances exceed the allowed values for clear sky. These exceptions indicate the presence of clouds and can be flagged as cloudy-sky measurements. Given that the MLS forward model may have a $\sigma_1 = \sim 3$ K uncertainty, the flagging criteria for high-tangent-height radiances will be set as

$$P - P_{110} > \sigma_1 \text{ for } \vec{p} < \vec{p}_1 \tag{6.1}$$

where $P_{110}$ is the modeled radiance for 110% saturation and $P$ is the observed radiance. $\vec{p}$ is the tangent pressure, while $\vec{p}_1$ is the cutoff pressure level for defining the high-tangent-height radiances, which is

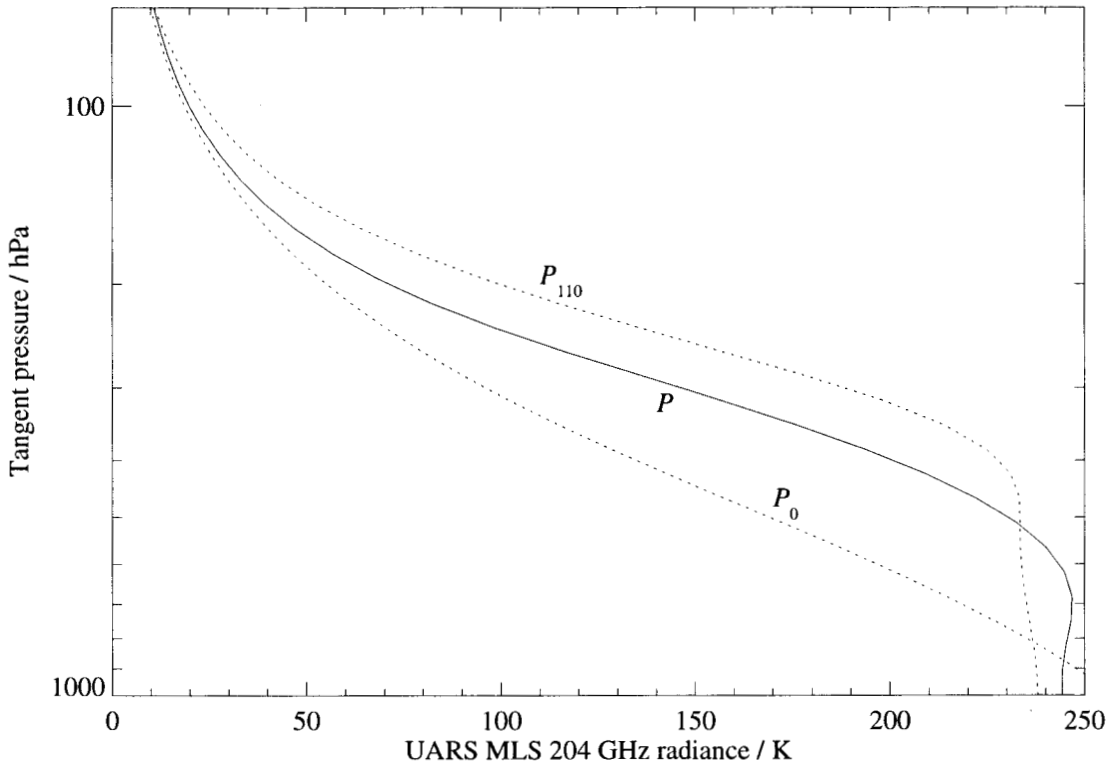



**Figure 6.1:** UARS MLS 204 GHz radiance measurement (solid line) and model calculations (dotted lines) for dry and for supersaturated moist air. The dry-air continuum is the major contributor to the dry profile, with small (3–5 K) contributions from minor constituents. The supersaturated profile is calculated assuming 110% relative humidity. The radiative transfer model was developed for UARS MLS upper-tropospheric humidity measurement (see Read et al. 1999), which uses dry-air and water vapor continuum coefficients determined empirically from the MLS data.

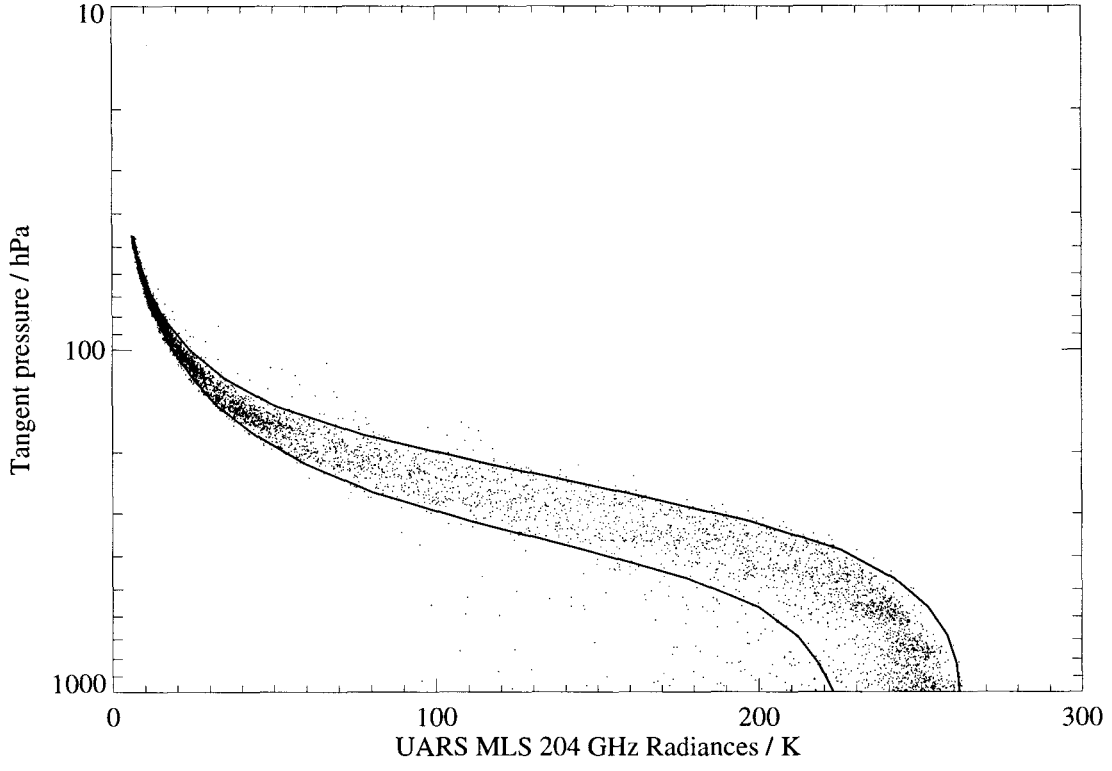



**Figure 6.2:** This plot shows observations of 204 GHz radiance from UARS MLS on 10 January 1992. The smooth lines approximately describe the theoretical radiance limits determined by the wet and dry continua in a cloud free atmosphere.

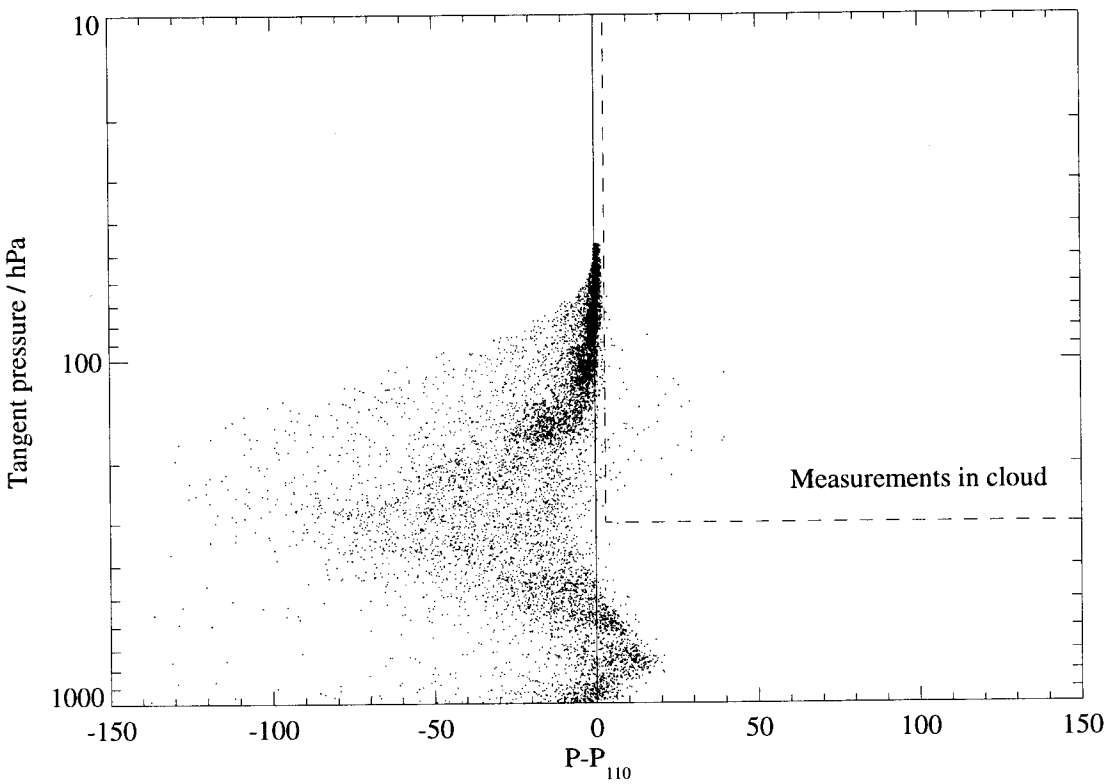



**Figure 6.3:** Radiance differences between UARS MLS observations (10 January 1992) and the maximum value allowed by the model for clear sky. Above ~300 hPa most differences are close to or less than 0 K except for those affected by clouds that show large positive values. The radiance maximum below ~8 km or ~300 hPa cannot be simply determined by $P_{110}$ because of complicated behaviors of the dry and wet profiles (see for example Figure 6.1).

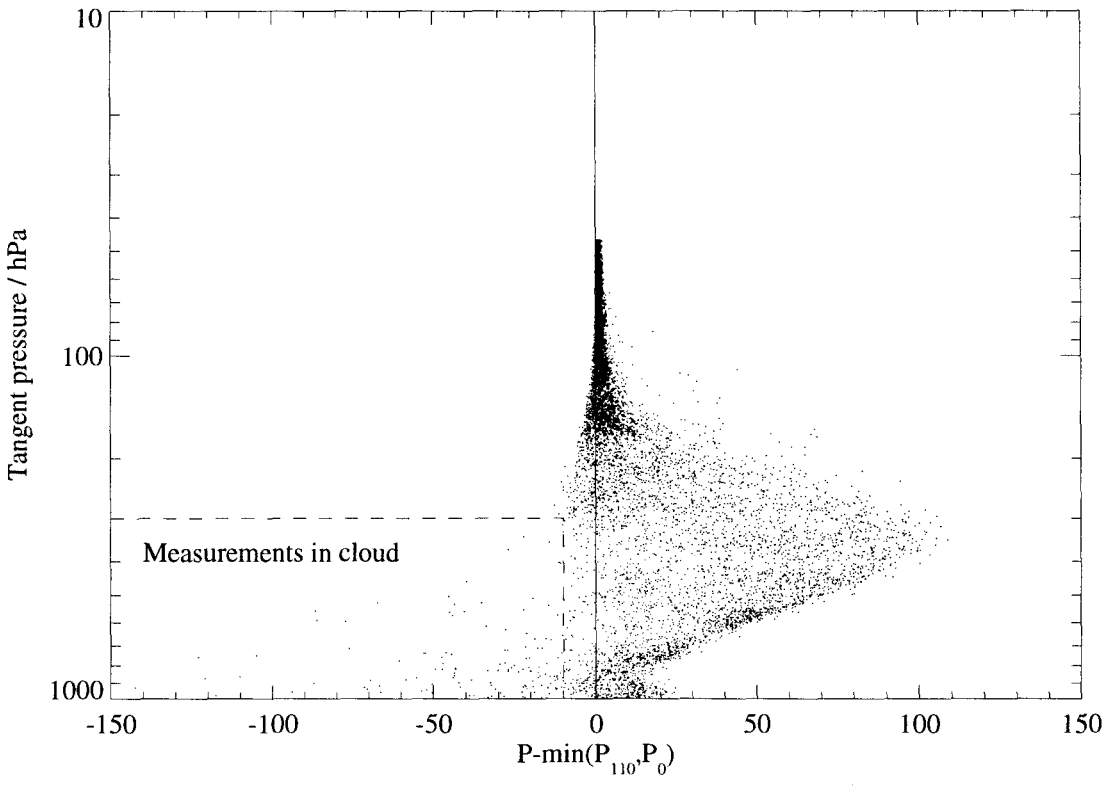



**Figure 6.4:** Radiance differences between 204 GHz UARS MLS 10 January 1992 measurements and the minimum values defined by the dry and saturated profiles. Most measurements are greater than the minimum values while a few are smaller. The negative differences, especially those less than -10 K, are likely to have been influenced by clouds. On the other hand, systematic error is also present showing a significant number of measurements in the 0 K and 10 K range above 316 hPa. This is probably caused by inaccurate pointing knowledge and constituent profiles (particularly, $O_3$ and $N_2O$).

316 hPa in the UARS case. The actual cutoff pressure and the forward model uncertainty (3 K in this case) will be the input parameters in the production software and will likely be frequency dependent.

Figure 6.4 illustrates the cloud flagging method for low altitude radiances. In this case, we check whether the measurement is below the low radiance threshold, mathematically

$$P - \min(P_{110}, P_0) < -\sigma_2 \text{ for } \vec{p} > \vec{p}_2 \tag{6.2}$$

where $P_0$ is the modeled radiance for dry air and $\vec{p}_2$ is the cutoff pressure level for this low-tangent-height criteria (316 hPa for UARS MLS). An uncertainty of $\sigma_2 = 10$ K is used for model and temperature error associated with UARS MLS 204 GHz, which is estimated from scatter in the plots. For EOS MLS this uncertainty will be provided by the forward model and will be frequency-dependent.

## 6.2 Flagging radiances contaminated by cloud

The MLS data processing algorithms will generate a number indicating the level of cloud contamination, called the *cloud induced radiance* (CIR). The CIR will simply be the deviation of the observed radiances from the clear-sky envelopes described above, with negative values indicating radiances below the lower limit of the envelope, and positive values indicating radiances larger than the upper limit.

The retrieval algorithm will generate CIRs for each radiometer and minor frame. These will then be compared with selected upper and lower thresholds (which can be a function of radiometer and tangent pressure). If a CIR exceeds the appropriate threshold then the corresponding radiances will be flagged and may not be used in the retrieval calculations.

The CIRs will be modified as the retrieval progresses, by comparing with the successively tighter envelopes obtained as the knowledge of the temperature profile and upper tropospheric humidity improves. The final CIR, essentially cloud signals, will be used for retrieving cloud parameters such as extinction coefficients and ice water content (IWC). In addition, the normal clear-sky retrieval will produce a cloud extinction coefficient (see Section 3.5.3). Since such extinction coefficient will account for all spectrally-flat radiances unexplained by the clear-sky model, it can be used to compare with CIRs determined with this procedure.

## 6.3 Uncertainties of cloud flags

The classification scheme described above depends crucially upon the accuracy of the MLS clear-sky forward model. Major uncertainty will arise from uncertainties in dry and wet continua. It's important to know these quantities to about 5% accuracy in order for the flagging method to work properly. In UARS MLS, the continua were obtained empirically from the data (Read et al. 1999), yielding a residual of $\sim 5-10\%$ at $10-20$ km tangent heights.

Good temperature and pressure measurements are required for this cloud flagging scheme. The CIR will be first calculated using temperature and tangent pressure obtained from a preliminary retrieval based on the $O_2$ radiances above $\sim 100$ hPa, and the scan model described in Section 3.4. Refinements of the CIR during the retrieval will be carried out as the retrieval process leads to improved estimates of these and other atmospheric species ($O_3$, $H_2O$, etc.)

## 6.4 Deriving cloud parameters from MLS observations

The final CIRs will go to a separate algorithm (in addition to MLS clear-sky retrieval) for cloud ice retrieval. The CIR to extinction coefficient and CIR to IWC relations will be derived using a full radiative transfer model that includes both absorption and scattering processes. This model is currently being developed, which will incorporate realistic parameterization for cloud ice particle size distributions. A detailed description of the model will be given in later versions of this document.

## 6.5 Summary of cloud flagging process

A summary of how the cloud flagging and characterization fit into the overall retrieval scheme is given in Figure 7.1 on page 52.

# Chapter 7
# Additional topics

This chapter considers some remaining issues with the MLS data processing algorithms, including topics which, while not strictly speaking part of the theoretical basis for the algorithms, are worthy of discussion.

## 7.1 Tuning of algorithms and strategy for post-launch operations

As discussed previously, in addition to the state of the atmosphere, several other parameters affect the radiances observed by EOS MLS. These parameters include spectroscopic data such as pressure broadening parameters, temperature exponents and pressure shift parameters; and instrumental parameters such as side band ratios, angular offsets between radiometers etc. While these parameters are measured as part of the pre-launch calibration, or taken from standard databases, it may be that better values for them can be obtained post-launch using the retrieval algorithms. These parameters can be included in the $\mathbf{x}^{\diamond}$ elements of the state vector as shown in Equation 4.1. Clearly the larger the chunks of data used in the retrieval algorithm, the more precise the estimates of the $\mathbf{x}^{\diamond}$ quantities. This technique can be particularly effective in cases where the same molecule is measured by MLS in multiple spectral bands. Given two sets of radiance measurements, and knowledge that the emission in each case is due to the same molecule, significantly more accurate estimates of the parameters such as spectroscopic terms may be obtained. Once 'optimum' values for these parameters are retrieved, it is possible to routinely produce 'definitive' measurements of the abundance of a molecule, being an optimum fit to all the relevant MLS radiance observations.

However, applying this technique immediately post launch is probably too ambitious, particularly as many of these parameters have a non-linear effect on the radiance observations. The best strategy is to process the data immediately post launch using somewhat cautious algorithms. As a separate process, the retrieval algorithm is run for large datasets to obtain optimal value of the $\mathbf{x}^{\diamond}$ quantities for use in later, more ambitious versions of the data processing software.

This section describes the details of this exercise, both in terms of the intended goals and the methods used to achieve them.

### 7.1.1 Composition from individual radiometers

The retrieval of parameters such as spectroscopy and sideband ratios in the $\mathbf{x}^{\diamond}$ part of the state vector allows for the retrieval of 'definitive' quantities such as the optimal ozone from all the MLS spectral bands. The behavior of these quantities will however be dependent on the operational mode of the MLS instrument. If part of the MLS instrument is turned off, for example the 240 GHz radiometer, these definitive products will be affected by the absence of some observations. For example, the definitive ozone will be reduced in quality due to the absence of `R3:240.B6F:O3` measurements.

The MLS instrument is designed to be operated in a 'power saving' mode if necessary, by switching off different radiometers for time scales of order a month. Such operations would make the use of these 'definitive' products unwise for studies involving long term timeseries and trend analyses.

For such analyses, it is more important that a temporally-consistent dataset is used, rather than the optimal one. For this reason it is intended that the MLS data processing algorithms, in addition to producing the definitive products from the combined observations in all the radiometers, will also routinely produce separate products describing the observations of atmospheric composition from each separate spectral band.

### 7.1.2 An example retrieval configuration

Figure 7.1 shows a possible implementation of the data processing flow which retrieves both the definitive and individual observations of the atmospheric observations. The goal of the first set of operations (those in the left hand column and the first one in the middle column) is to obtain an optimal value of temperature and upper tropospheric and lower stratospheric water vapor, along with the cloud information from the 118 GHz oxygen line and a subset of the $H_2O$ radiance information.

Following this, simple 1D retrievals of each constituent are performed separately to obtain a suitable starting point for future phases. At this point the algorithm splits, the operations in the center column are designed to obtain the 'optimal' products from all the radiance information. The second set of operations, in the far right column retrieve the separate composition measurements from each spectral band.

## 7.2 Quality control, exception handling and related issues

### 7.2.1 Quality of retrieved data

In addition to retrieving an optimum state vector, the MLS data processing algorithms will compute an estimated uncertainty for each element of the state vector. As described in Section 4.3.3, this uncertainty will usually be taken from the diagonal elements of the inverse of the $\left[\mathbf{S_a} + \sum_i \mathbf{K}_i^T \mathbf{S}_i^{-1} \mathbf{K}_i\right]$ matrix used in the preconditioner calculation, although other sources of uncertainty information are possible including the full covariance matrix for the solution state vector.

As described in Section 3.2.3, the uncertainty on the retrieved result should always be compared with the uncertainty given on the *a priori* information. This comparison, along with the uncertainty information itself will form a major part of quality control. In the UARS MLS case, the uncertainty is set negative if it is greater than half of the *a priori* uncertainty. This serves as a useful flag to the users of the data, to indicate where data should be approached with caution. Similar flags may be implemented for EOS MLS, but these issues remain to be decided.

Another source of quality control information will be the $\chi^2$ information for the radiance measurements. Cases where the retrieval has converged on an inappropriate solution, or where the radiances are poor for some reason, will be clearly indicated by a high value of $\chi^2$. A complete set of $\chi^2$ statistics will be produced by the data processing algorithms, giving the values of $\chi^2$ for each major frames worth of radiance observations for each band. This $\chi^2$ information will also form the basis of a simple quality flag for each product, indicating the validity of the data, as was done for UARS MLS.

### 7.2.2 Bad or missing radiances

If a radiance observation is missing or marked bad for whatever reason, it is simply not included in the retrieval calculation (in the same manner as those radiances marked as contaminated by cloud as described in Section 6.2.) If several consecutive radiances are missing, such as a whole major frame's radiances for a band, retrieval is still possible, as the *a priori* information, along with information from

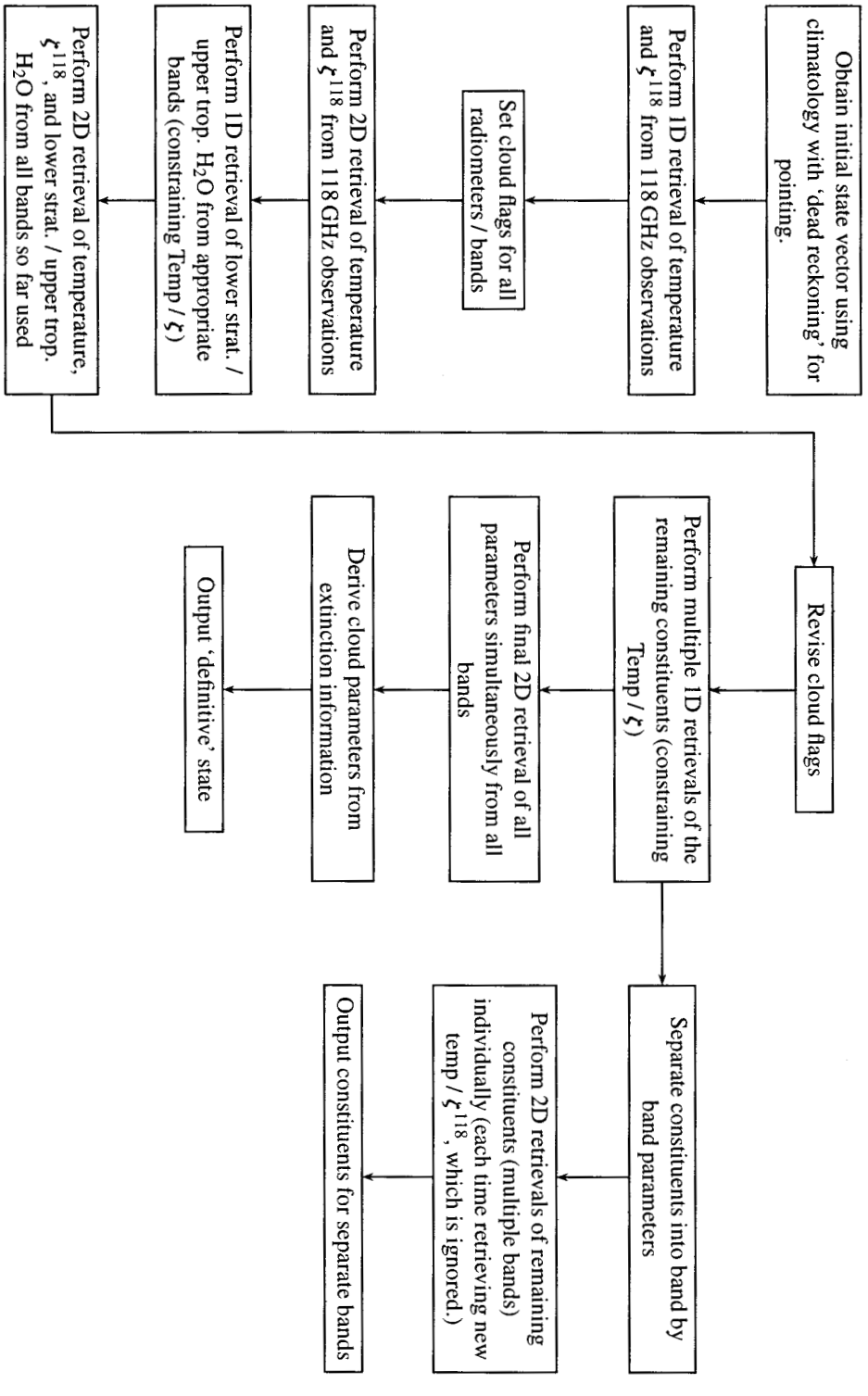



**Figure 7.1:** An example configuration for the EOS MLS Level 2 processing scheme.

the adjacent scans, will influence the retrieval for the corresponding profile. The retrieved uncertainties however would reflect the comparative lack of information for the corresponding profile.

If several consecutive major frames worth of radiances were missing, the retrieval algorithm will take this into account when dividing the dataset into chunks, using the boundaries of the region of missing data as the edges of the chunks.

### 7.2.3 Numerical exceptions

The retrieval calculations described here are sufficiently well posed and numerically stable (when scaled as described in Section 4.5) that occurrences such as division by zero, or requesting the square root of a negative number should never occur. For this reason, no special handling is needed for such events; any attempt to perform such a calculation will be indicative of a 'bug' in the program, and so should simply bring the processing to an immediate halt with an appropriate error message.

## 7.3 Suitability of the algorithm to modern computer architectures

The size of the EOS MLS data processing task is such that a parallel computer system will be required for the retrieval calculations. Retrieval calculations have in the past been described as 'embarrassingly parallel' problems. There are many different ways one can choose to break up the problem into independent tasks; one could divide the task up by profile, surface, spectral band, channel etc.

The proposed algorithm retrieves the data in chunks of interrelated profiles, so the task is not quite so easy to divide up; however, the task is still relatively simple. Section 4.4.1 showed that the dominant calculations of the algorithm scale as $N$ where $N$ is the size of the chunk. The limiting factor on the value of $N$ is clearly thus going to be the memory capacity of the computer system(s) used. The value of $N$ should be chosen such that all the $\mathbf{K}$ and $\mathbf{K}^T\mathbf{S}_y^{-1}\mathbf{K}$ matrices fit in memory. The use of memory paging (swapping memory out to disk) should be avoided as this would dramatically slow down the speed of the computations.

The use of chunks can also lead to efficiency within the forward model calculations. Given a set of $N$ scans for which to compute forward model radiances, one can gain efficiency by doing all the calculations for one spectral band together. In this manner one could for example load all the relevant spectroscopic information for band 1, compute all $N$ forward models for this band, then go onto band 2. The alternative course, whereby all the forward models for one scan are calculated together is less efficient as the computer is constantly switching spectroscopy databases.

Perhaps the simplest parallel machine is a network of workstation type computers; if these were the machines available, then one manner in which to perform the retrieval calculation would be to assign each chunk of the dataset to a different processor. Clearly each computer would need a large amount of memory in order that an appropriately large value of $N$ can be obtained. An alternative to this approach is to divide the chunks into smaller groupings of, say, five profiles, and distribute these among the available processors. The processors would then collaborate by performing all the calculations relevant to their elements of the chunk, using message passing to communicate to their 'neighabours' any components of the matrices they require for their calculations.

Other parallel machines work on either the *symmetric multi processing* (SMP) or the *massively parallel processing* (MPP) models. The distinction between these is becoming somewhat blurred; for example, the current MLS Science Computing Facility (SCF) compute server is a Silicon Graphics Origin 2000 machine, this has an MPP like architecture, but appears to the programmer to be an SMP type machine. These types of parallel machines are the prefered architecture for the EOS MLS retrieval algorithms. On these machines the different processors can work together much more coherently than

in the networked workstations model, due to the fast communications that are possible between them.

In this model the EOS MLS retrieval algorithm would run sequentially through the chunks, running the forward models for the whole chunk one band at a time as described above. The forward model calculations could be parallelized by channel, profile etc. in a manner to be determined. The matrix computations involved in the retrieval calculations are fairly easy to parallelize, many standard methods for this exist, (see Golub and VanLoan 1996, for example.)

## 7.4 Computational requirements

Section 4.4.1 outlined the number of floating point operations involved in the various parts of the retrieval calculation. As explained in that section, the number of floating point operations can only serve as a rule of thumb for the expected amount of computation time required. Table 4.1 showed that the most intensive calculation in the algorithm is the computation of the $\mathbf{K}^T\mathbf{S}^{-1}\mathbf{K}$ matrix, which takes $Np^2n^2m$ operations. As explained previous, this is an approximate figure, as it does not take into account sparsity within the sub matrices, or the lack of sensitivity of say the `R1:118` observations to quantities such as $O_3$ and ClO.

Clearly, the more complex the retrieval system, the larger the size of the computational task. The most complex retrieval system in the MLS case will be the retrieval of all the 'definitive' products from all the GHz radiance observations (the THz observations are likely to be considered separately in the routine processing.) We will consider a retrieval of such a system for the case where the data are retrieved in chunks of 70 profiles each (60 profiles is 1/4 of an orbit, 5 profile overlaps at each end gives $N = 70$.) Only the standard 25-channel filter bank measurements are used. The complete list of all the species considered and bands used is shown in Table 7.1.

If all the radiances are used, then the $\mathbf{K}_i$ matrices occupy 14.5 Gigabytes (Gb). If only 20% of the radiances are used (as described in Section 4.4.5) this size reduces to 2.9 Gb. In either case the $\sum_i \mathbf{K}_i^T\mathbf{S}_i^{-1}\mathbf{K}$ matrix occupies 0.8 Gb. The computation of this matrix takes $2.36\times10^{12}$ operations if all the radiances are used, and $4.73\times10^{11}$ if only 20% of the radiances are used. These correspond to execution times of 39.4 and 7.8 minutes respectively when running at perfect efficiency on a 1 Gflop ($1\times10^9$ floating point operations per second) computer. This is comparable to the ~25 minutes it takes MLS to measure the 60 profiles in the main part of the chunk. Given the additional computation time required to perform the forward model calculations and the earlier phases of the retrieval, it is clear that processing the data at a rate comparable to real time is perfectly feasible given the current levels of computing power (note that the current MLS Science Computing Facility compute server has been bench-marked at over 4 Gflops.)

### 7.4.1 Re-blocking of matrices

In large dense matrix calculations typically the best efficiency is achieved by dividing the matrices involved into blocks and dealing with the calculation one block at at time. This is the method adopted by libraries such as LAPACK and BLAS. This method is clearly similar to the one proposed here, except that in our case many of the blocks are zero. In the case of the dense matrix calculations the size of the blocks is chosen to best suit the computer used (by comparison to the size of the cache, and length of the floating point pipelines etc.). The size of the blocks in the MLS calculation is determined by the measurement system.

It is possible that the efficiency of the calculations could be improved by re-dividing the matrices involved into larger blocks. While this would involve storing and considering more zeros, efficiency could still be gained if the size of the new submatrices was conducive to greater efficiency. Such a

re-blocking exercise could be performed if necessary after the forward model calculation has assembled the **K** matrices.

## 7.5 Data volumes

The volume of geophysical data produced by the MLS retrieval algorithms is small compared to the amount of radiance data. The EOS MLS overview document (Waters 1999) describes the volumes of data required in tables 8.8 and 8.9, with a total data volume of about 55 Mb per day of data.

In addition to geophysical parameters with their error estimates, the retrieval algorithm will also need to routinely output state vector elements such as $\zeta^{118}$ etc., and statistics such as $\chi^2$ values, this will add about 10 Mb to the data volume. The algorithm will also be capable of outputting quantities such as full matrices, and forward model radiance estimates, though the former of these will probably not be produced in the routine data processing, due to the large data volume involved.

## 7.6 Validation of Level 2 data products and algorithms

The approach to the validation of the EOS MLS data follows the procedure used successfully for UARS MLS, and is summarized in the overview document (Waters 1999). Some validation of the Level 2 algorithm itself will also be required. A vital tool for this validation is the use of simulated instrument data sets (SIDS). Given a model of the state of the atmosphere, it is possible, using the forward model calculation to generate a set of simulated radiances describing the observations MLS would make were the atmosphere in the modeled state. Retrieval calculations using these radiances both in the presence and absence of instrumental noise and systematic errors yield valuable insight into the performance of the retrieval algorithm.

In addition to such tests, the individual components of the retrieval algorithm will need to be validated. In most cases this validation method is clear; for example, the preconditioned conjugate gradient solver algorithm can be tested on a known system, or by comparison with a complete Cholesky Decomposition based matrix solving algorithm.

## 7.7 Alternative methods considered

The retrieval algorithm proposed here is by no means the only reasonable algorithm that could be implemented for the EOS MLS retrieval problem. This section outlines some of the alternative methods considered. All the methods considered use optimal estimation, given the obvious superiority of the optimal methods for these problems (for more discussion of this see Rodgers 1976.)

### 7.7.1 A moving state vector

An alternative to retrieving data in chunks is to construct a state vector consisting of a set of ~5 profiles, then using a sequential method to assimilate the radiance observations, either one radiance or one scan at a time, moving the state vector through the dataset one profile at a time. The disadvantage of such a scheme is that forward and backward passes would probably be required for accurate results. The results from these would have to be merged by their full covariance matrices, a time consuming and memory intensive calculation. In order to reduce the memory requirements, it would be necessary to break the day into chunks as in the proposed algorithm.

**Table 7.1:** This table shows the configuration of the retrieval for which the number of operations required has been computed. For each state vector component, the range and resolution of the data are given, along with a matrix indicating whether derivative information is stored for each 25 channel filter bank in the GHz module.

| Quantity | Range | Resolution | R1A:118.B1F:PT | R2:190.B2F:H2O | R2:190.B3F:N2O | R2:190.B4F:HNO3 | R2:190.B5F:CLO | R2:190.B6F:O3 | R3:240.B7F:O3 | R3:240.B8F:PT | R4:640.B9F:CO | R4:640.B10F:CLO | R4:640.B11F:BRO | R4:640.B12F:N2O | R4:640.B13F:HCL | R4:640.B14F:O3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ptan118 | N/A | 120 values | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Temperature | 1000.0 – 0.01 hPa | 6 / decade | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $O_3$ | 1000.0 – 0.01 hPa | 6 / decade | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | | | | ✓ | ✓ |
| $H_2O$ | 1000.0 – 0.01 hPa | 6 / decade | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| $N_2O$ | 1000.0 – 1.0 hPa | 6 / decade | | ✓ | ✓ | ✓ | | | | | | ✓ | ✓ | ✓ | | |
| $HNO_3$ | 1000.0 – 1.0 hPa | 6 / decade | | ✓ | ✓ | ✓ | | | | | | | | | | |
| ClO | 1000.0 – 1.0 hPa | 6 / decade | | | | | ✓ | | | | | ✓ | ✓ | ✓ | | |
| CO | 1000.0 – 0.01 hPa | 6 / decade | | | | | | | | ✓ | | | | | | |
| HCl | 1000.0 – 0.1 hPa | 6 / decade | | | | | | | | | | | | | ✓ | ✓ |
| R1 extinction | 1000.0 – 0.01 hPa | 6 / decade | ✓ | | | | | | | | | | | | | |
| R2 extinction | 1000.0 – 0.01 hPa | 6 / decade | | ✓ | ✓ | ✓ | ✓ | ✓ | | | | | | | | |
| R3 extinction | 1000.0 – 0.01 hPa | 6 / decade | | | | | | | ✓ | ✓ | ✓ | | | | | |
| R4 extinction | 1000.0 – 0.01 hPa | 6 / decade | | | | | | | | | | ✓ | ✓ | ✓ | ✓ | ✓ |

# Appendix A
# Algorithms for other MLS products

## A.1 Tropopause pressure

One of the diagnostic quantities produced by the MLS Level 2 processing will be the pressure at the tropopause. This quantity will be very useful in many dynamical studies undertaken using MLS data, particularly those involving stratosphere/troposphere exchange. This quantity will typically be derived by fitting a function such as a cubic spline to the retrieved MLS temperature profile in the region around the tropopause, and inferring the pressure of the coldest region. It is also intended that the algorithms will produce a tropopause pressure estimate as determined from the correlative data sources such as National Center for Environmental Prediction (NCEP), or the GSFC Data Assimilation Office (DAO).

A similar quantity that may be produced is the pressure at the *hydropause*. This is the pressure at which the water vapor profile has its minimum. This will be obtained in a similar fasion to the tropopause pressure, by fitting some appropriate function to the profile, and obtaining the pressure at the minimum value.

## A.2 Column products

In addition to providing profiles of atmospheric species, the Level 2 processing will also output stratospheric column abundances of selected species (e.g. ozone.) These are obtained by integrating the abundance profiles from the tropopause (as obtained above) to the top of the retrieval range, and then converting the product into appropriate units (e.g. Dobson units).

The full details of the calculations can be found in the next section, which is a modified reproduction of an earlier document by W.G. Read and J.W. Waters.

## A.3 Column abundances of MLS profiles.

This is to document expressions for column abundances in the vertical profiles retrieved from MLS.

Let $N(z_1, z_2)$ be the vertical column of molecules (per square meter) between heights $z_1$ and $z_2$. This is given by:

$$N(z_1, z_2) \;=\; \int_{z_1}^{z_2} f(z)\, n(z)\, \mathrm{d}z, \tag{A.1}$$

where $f(z)$ is the volume mixing ratio at height $z$ of the species being considered, and $n(z)$ is the *total* ('air') number density at $z$.

We convert to pressure $p$, which is the vertical coordinate for MLS retrievals, by using hydrostatic equilibrium:

$$\mathrm{d}p(z) \;=\; -\rho(z)\, g(z)\, \mathrm{d}z, \tag{A.2}$$

### 7.7.2 A 'sequential' approach

Another approach considered uses the same chunk-based state vector as the proposed algorithm. However, in this case the radiances are inserted sequentially. In some respects this is similar to the moving state vector method above; however, in this case the measurements are assimilated from the higher tangent points downwards. This is similar to the UARS MLS version 4 data processing, and related to the 'onion peeling' technique. By inserting measurements from the top down, the two dimensional aspect of the problem is dealt with explicitly, as the knowledge of the state of a particular profile / surface is retrieved before it is required in forward model calculations for lower tangent points (consideration Figure 3.6 should make this clearer.)

The disadvantage of this technique is that the forward model computations are required in an inefficient order. Because the forward model involves a convolution with an instrumental field of view, the radiative transfer must be computed for a large vertical range of tangent heights, even if only one tangent height is required. In addition, this method is more dependent on the quality of the retrieved data at high altitudes, as they are used in the retrieval of lower altitude data, these high altitude observations are typically poor, due to the low radiance values (and thus low signal to noise ratios) involved.

where $\rho(z)$ is the mass density of air at $z$ and $g(z)$ is the gravitational acceleration. Converting from mass density $\rho$ to number density $n$, and neglecting the small variation[1] in $g$, (A.2) becomes

$$\mathrm{d}p(z) \;=\; -\frac{M\,g}{A}\,n(z)\;\mathrm{d}z, \tag{A.3}$$

where $M$ is the 'effective' molecular weight of air and $A$ = Avogadro's number ($6.022\times10^{23}$ molecules/mole). Using $n(z)\,\mathrm{d}z$ from (A.3) in (A.1) gives

$$N(z_1, z_2) = N(p_2, p_1) \;=\; \frac{A}{Mg}\int_{p_2}^{p_1} f(p)\;\mathrm{d}p, \tag{A.4}$$

where $p_1$ is the pressure at $z_1$ and $p_2$ is the pressure at $z_2$, and it is assumed that $p_1 > p_2$.

If the mixing ratio $f(p)$ is a constant between $p_1$ and $p_2$ then equation (A.4) becomes

$$N(p_2, p_1) \;=\; \frac{A}{Mg}\,f\,\Delta p \tag{A.5}$$

$$\;=\; \frac{(6.022\times10^{23}\ \text{molecules/mole})}{(28.97\ \text{g/mole})(9.71\ \text{ms}^{-2})}\,f\,\left(\Delta p \text{kg m}^{-1}\,\text{s}^{-2}\right) \tag{A.6}$$

$$\;=\; 2.14\times10^{24}\,f\,\Delta p\ \text{molecules}/m^2\ \text{(for } p \text{ in Pascals)} \tag{A.7}$$

$$\;=\; 2.14\times10^{26}\,f\,\Delta p\ \text{molecules}/m^2\ \text{(for } p \text{ in hPa)}, \tag{A.8}$$

where

$$\Delta p \;\stackrel{\text{def}}{=}\; p_1 - p_2, \tag{A.9}$$

$g = 9.71\ \text{m/s}^2$ (representative of 45° latitude and 30 km altitude) and 28.97 g/mole for the molecular weight of air[2] have been used in (A.6). For $f = 1$ we obtain the vertical column of 'air' and (A.8) shows that an air layer of 'thickness' 1 hPa contains $2.14\times10^{26}$ molecules per square meter.[3]

The mixing ratio profiles retrieved from MLS are given by

$$f(p) \;=\; \sum_j f_j\,\eta_j(p), \tag{A.10}$$

where $\eta_j(p)$ are the 'basis functions' used for representing the profile and $f_j$ are retrieved coefficients. Putting (A.10) in (A.4) gives

$$N(p_1, p_2) \;=\; \frac{A}{Mg}\int_{p_2}^{p_1}\sum_j f_j\,\eta_j(p)\;\mathrm{d}p \tag{A.11}$$

$$\;=\; \frac{A}{Mg}\sum_j f_j\int_{p_2}^{p_1}\eta_j(p)\;\mathrm{d}p. \tag{A.12}$$

The basis functions currently used for MLS retrievals are triangular in log $p$ and are sketched in figure A.1.

[1]The effective gravitational acceleration at Earth's surface varies 0.5% between 9.780 m/s$^2$ at the equator and 9.832 m/s$^2$ at the pole, and is 9.806 m/s$^2$ at 45° latitude. It also decreases 1% for each 32 km increase in height, for heights much less than Earth's radius. See, for example, Fleagle and Businger 1963 and Hess 1959.

[2]See, for example, Hess 1959.

[3]Note that this is independent of temperature, which determines the *height* thickness of the layer (through the gas law).

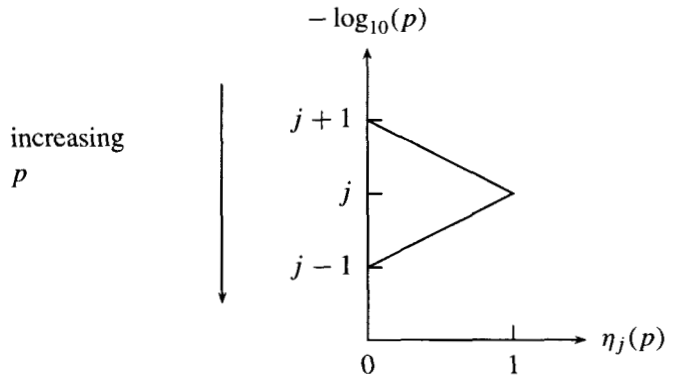



**Figure A.1:** The MLS basis functions.

The mathematical expressions for these basis functions are

$$\eta_j(p) = 0 \quad \text{for } p < p_{j+1} \tag{A.13}$$

$$\eta_j(p) = (\log_{10} p - \log_{10} p_{j+1})/\Delta_{j+}^{\log_{10} p} \quad \text{for } p_j \geq p \geq p_{j+1} \tag{A.14}$$

$$\eta_j(p) = (\log_{10} p_{j-1} - \log_{10} p)/\Delta_{j-}^{\log_{10} p} \quad \text{for } p_{j-1} \geq p \geq p_j \tag{A.15}$$

$$\eta_j(p) = 0 \quad \text{for } p > p_{j-1}, \tag{A.16}$$

where

$$\Delta_{j+}^{\log_{10} p} = \log_{10} p_j - \log_{10} p_{j+1} \tag{A.17}$$

$$\Delta_{j-}^{\log_{10} p} = \log_{10} p_{j-1} - \log_{10} p_j, \tag{A.18}$$

Using (A.13)–(A.18), the integral in (A.12) can be evaluated. It is broken up into that for the 'bottom' portion of the basis function ranging between $j-1$ and $j$ (indicated by superscript '$-$'), and the 'top' portion ranging between $j$ and $j+1$ (indicated by superscript '$+$'). We assume that the integration limits $p_1$, $p_2$ occur *outside* the region where $\eta_j(p)$ is non-zero.

First, for integration over the the 'bottom' portion of $\eta_j(p)$:

$$I_j^- \stackrel{\text{def}}{=} \int_{p_j}^{p_{j-1}} \eta_j(p)\, \mathrm{d}p \tag{A.19}$$

$$= \frac{1}{\Delta_{j-}^{\log_{10} p}} \int_{p_j}^{p_{j-1}} \left(\log_{10} p_{j-1} - \log_{10} p\right)\, \mathrm{d}p \tag{A.20}$$

$$= \frac{1}{\Delta_{j-}^{\log_{10} p}\, \ln 10} \left\{ \ln p_{j-1} \int_{p_j}^{p_{j-1}} \mathrm{d}p - \int_{p_j}^{p_{j-1}} \ln p\, \mathrm{d}p \right\} \tag{A.21}$$

$$= \frac{1}{\Delta_{j-}^{\log_{10} p}\, \ln 10} \left\{ \left(p_{j-1} - p_j\right) \ln p_{j-1} - \left(p \ln p - p\right) \Big|_{p_j}^{p_{j-1}} \right\} \tag{A.22}$$

$$= \frac{1}{\Delta_{j-}^{\log_{10} p}\, \ln 10} \left\{ p_{j-1} \ln p_{j-1} - p_j \ln p_{j-1} - p_{j-1} \ln p_{j-1} + p_{j-1} + p_j \ln p_j - p_j \right\} \tag{A.23}$$

$$= \frac{1}{\Delta_{j-}^{\log_{10} p}\, \ln 10} \left\{ -p_j \ln p_{j-1} + p_{j-1} + p_j \ln p_j - p_j \right\} \tag{A.24}$$

$$= -p_j \left\{ \frac{\log_{10} p_{j-1} - \log_{10} p_j}{\Delta_{j-}^{\log_{10} p}} \right\} + \frac{p_{j-1} - p_j}{\Delta_{j-}^{\log_{10} p}\, \ln 10} \tag{A.25}$$

$$= -p_j + \frac{1}{\ln 10} \frac{\Delta_{j-}^{p}}{\Delta_{j-}^{\log_{10} p}} \tag{A.26}$$

$$= -p_j + \frac{\Delta_{j-}^{p}}{\Delta_{j-}^{\ln p}}, \tag{A.27}$$

where

$$\Delta_{j-}^{p} \stackrel{\text{def}}{=} p_{j-1} - p_j \tag{A.28}$$

$$\Delta_{j-}^{\ln p} \stackrel{\text{def}}{=} \ln p_{j-1} - \ln p_j. \tag{A.29}$$

(A.18) has been used in (A.26), and

$$\int \ln p\, \mathrm{d}p = p \ln p - p \tag{A.30}$$

has been used in (A.22), and

$$\log_{10} p = \frac{\ln p}{\ln 10} \tag{A.31}$$

has been used in (A.21,A.25,A.27).

Now integration over the the 'top' portion of $\eta_j(p)$ is performed in a similar fashion:

$$I_j^+ \stackrel{\text{def}}{=} \int_{p_{j+1}}^{p_j} \eta_j(p)\, \mathrm{d}p \tag{A.32}$$

$$= \frac{1}{\Delta_{j+}^{\log_{10} p}} \int_{p_{j+1}}^{p_j} \left(\log_{10} p - \log_{10} p_{j+1}\right)\, \mathrm{d}p \tag{A.33}$$

$$= \frac{1}{\Delta_{j+}^{\log_{10} p} \ln 10} \left\{ \int_{p_{j+1}}^{p_j} \ln p\, \mathrm{d}p - \ln p_{j+1} \int_{p_{j+1}}^{p_j} \mathrm{d}p \right\} \tag{A.34}$$

$$= \frac{1}{\Delta_{j+}^{\log_{10} p} \ln 10} \left\{ (p \ln p - p) \Big|_{p_{j+1}}^{p_j} - \left(p_j - p_{j+1}\right) \ln p_{j+1} \right\} \tag{A.35}$$

$$= \frac{1}{\Delta_{j+}^{\log_{10} p} \ln 10} \left\{ p_j \ln p_j - p_j - p_{j+1} \ln p_{j+1} + p_{j+1} - p_j \ln p_{j+1} + p_{j+1} \ln p_{j+1} \right\} \tag{A.36}$$

$$= \frac{1}{\Delta_{j+}^{\log_{10} p} \ln 10} \left\{ p_j \ln p_j - p_j + p_{j+1} - p_j \ln p_{j+1} \right\} \tag{A.37}$$

$$= p_j \left\{ \frac{\log_{10} p_j - \log_{10} p_{j+1}}{\Delta_{j+}^{\log_{10} p}} \right\} - \frac{p_j - p_{j+1}}{\Delta_{j+}^{\log_{10} p} \ln 10} \tag{A.38}$$

$$= p_j - \frac{1}{\ln 10} \frac{\Delta_{j+}^{p}}{\Delta_{j+}^{\log_{10} p}} \tag{A.39}$$

$$= p_j - \frac{\Delta_{j+}^{p}}{\Delta_{j+}^{\ln p}}, \tag{A.40}$$

where

$$\Delta_{j+}^{p} \stackrel{\text{def}}{=} p_{j+1} - p_j \tag{A.41}$$

$$\Delta_{j+}^{\ln p} \stackrel{\text{def}}{=} \ln p_j - \ln p_{j+1}. \tag{A.42}$$

The vertical column $N_j$ represented by a *single* retrieval coefficient $f_j$ is given by the $j^{\text{th}}$ term in (A.12), with the integral evaluated between $j+1$ and $j-1$:

$$N_j = \frac{A}{Mg} f_j \int_{p_{j+1}}^{p_{j-1}} \eta_j(p)\, \mathrm{d}p \tag{A.43}$$

$$= \frac{A}{Mg} f_j \left[ I_j^- + I_j^+ \right] \tag{A.44}$$

$$= \frac{A}{Mg} f_j \left[ \frac{\Delta_{j-}^{p}}{\Delta_{j-}^{\ln p}} - \frac{\Delta_{j+}^{p}}{\Delta_{j+}^{\ln p}} \right]. \tag{A.45}$$

The EOS MLS retrievals will use basis functions having

$$\Delta_{j+}^{\log_{10} p} = \Delta_{j-}^{\log_{10} p} = \frac{1}{12} \tag{A.46}$$

for all $j$. This gives

$$\Delta_{j+}^{\ln p} = \Delta_{j-}^{\ln p} = \frac{1}{12} \ln 10 = 0.1919 \tag{A.47}$$

and

$$\Delta^p_{j-} = \left(10^{1/12} - 1\right) p_j = 0.2115\, p_j \quad \text{(A.48)}$$

$$\Delta^p_{j+} = \left(1 - 10^{-1/12}\right) p_j = 0.1746\, p_j. \quad \text{(A.49)}$$

The column represented by the lower portion of the basis function is then

$$N_j^- = \frac{A}{Mg} f_j I_j^- \quad \text{(A.50)}$$

$$= \frac{A}{Mg}\left[\frac{12}{\ln 10}\left(10^{1/12} - 1\right) - 1\right] f_j\, p_j \quad \text{(A.51)}$$

$$= 0.102\, \frac{A}{Mg} f_j\, p_j \quad \text{(A.52)}$$

$$= 2.18\times 10^{25}\, f_j\, p_j \text{molecules/m}^2 \text{ (for } p_j \text{ in hPa)}, \quad \text{(A.53)}$$

and the column represented by the upper portion is

$$N_j^+ = \frac{A}{Mg} f_j I_j^+ \quad \text{(A.54)}$$

$$= \frac{A}{Mg}\left[1 - \frac{12}{\ln 10}\left(1 - 10^{-1/12}\right)\right] f_j\, p_j \quad \text{(A.55)}$$

$$= 0.090\, \frac{A}{Mg} f_j\, p_j \quad \text{(A.56)}$$

$$= 1.92\times 10^{25}\, f_j\, p_j \text{molecules/m}^2 \text{ (for } p_j \text{ in hPa)} \quad \text{(A.57)}$$

The column represented by the complete basis function, (A.45), becomes

$$N_j = \frac{A}{Mg}\frac{12}{\ln 10}\left(10^{1/12} + 10^{-1/12} - 2\right) f_j\, p_j \quad \text{(A.58)}$$

$$= 0.192\, \frac{A}{Mg} f_j\, p_j \quad \text{(A.59)}$$

$$= 4.12\times 10^{25}\, f_j\, p_j \text{molecules/m}^2 \text{ (for } p_j \text{ in hPa)}. \quad \text{(A.60)}$$

53% of the column represented by the basis function is in the lower portion and 47% is in the upper portion.

The total vertical column represented in the currently-retrieved MLS profiles is then

$$N = 4.12\times 10^{25} \sum_j f_j\, p_j \text{molecules/m}^2 \text{ (for } p_j \text{ in hPa)}. \quad \text{(A.61)}$$

For ozone it is convenient to express the column in Dobson Units (1 DU = $2.687\times 10^{20}$ molecules/m$^2$) and the mixing ratio in ppmv (1 ppmv = $1\times 10^{-6}$). Equation (A.61) then becomes

$$\text{DU} = 0.153 \sum_j f_j\, p_j \text{(for } p_j \text{ in hPa and } f_j \text{ in ppmv)}. \quad \text{(A.62)}$$

The column *between* retrieval levels $i$ and $u$ is

$$N(p_i, p_u) = N_i^+ + N_u^- + \sum_{j=i+1}^{j=u-1} N_j. \quad \text{(A.63)}$$

If $f_j$ is a constant for all $j$, it can be shown that (A.63) reduces to (A.5).

The column between *adjacent* retrieval levels $j$ and $j+1$ is

$$N(p_j, p_{j+1}) = N_j^+ + N_{j+1}^- \tag{A.64}$$

$$= \frac{A}{Mg}\left[1 - \frac{12}{\ln 10}\left(1 - 10^{-1/12}\right)\right] f_j\, p_j + \frac{A}{Mg}\left[\frac{12}{\ln 10}\left(10^{1/12} - 1\right) - 1\right] f_{j+1}\, p_{j+1} \tag{A.65}$$

$$= \frac{A}{Mg}\left[1 - \frac{12}{\ln 10}\left(1 - 10^{-1/12}\right)\right] f_j\, p_j + \frac{A}{Mg}\left[\frac{12}{\ln 10}\left(1 - 10^{-1/12}\right) - 10^{-1/12}\right] f_{j+1}\, p_j \tag{A.66}$$

$$= \frac{A}{Mg}\left[0.090\, f_j + 0.084\, f_{j+1}\right] p_j, \tag{A.67}$$

where $p_{j+1} = 10^{-1/12}\, p_j$ has been used in (A.66).

It is interesting to compare (A.67) with the approximate expression which computes the column between two levels by assuming a constant mixing ratio equal to the average of retrieved values at $j$ and $j+1$. This approximate expression is

$$N(p_j, p_{j+1}) \approx \frac{A}{Mg}\left[0.5\, f_j + 0.5\, f_{j+1}\right]\left[p_j - p_{j+1}\right] \tag{A.68}$$

$$\approx \frac{A}{Mg}\left[0.5\, f_j + 0.5\, f_{j+1}\right]\left[1 - 10^{-1/12}\right] p_j \tag{A.69}$$

$$\approx \frac{A}{Mg}\left[0.087\, f_j + 0.087\, f_{j+1}\right] p_j. \tag{A.70}$$

If the mixing ratio *is* constant between $j$ and $j+1$, then (A.70) and (A.67) give the same answer — as they should. The error in (A.70) depends upon the the difference in $f_j$ and $f_{j+1}$; worst-case error, when either $f_j$ or $f_{j+1}$ is zero, is approximately 10%.

Table A.1 summarizes expressions useful in calculating column abundances in the profiles retrieved from MLS.

**Table A.1:** Some useful expressions in calculating column abundances from retrieved MLS profiles.

| Column abundance | Molecules/m$^2$ (note 1) | Dobson Units (note 2) |
|---|---|---|
| in lower portion of $\eta_j(p)$ | $2.18\times10^{25}\, f_j\, p_j$ | $0.071\, f_j\, p_j$ |
| in upper portion of $\eta_j(p)$ | $1.92\times10^{25}\, f_j\, p_j$ | $0.080\, f_j\, p_j$ |
| in all of $\eta_j(p)$ | $4.12\times10^{25}\, f_j\, p_j$ | $0.153\, f_j\, p_j$ |
| between levels $j$ and $j+1$ | $\left(1.92 f_j + 1.80 f_{j+1}\right)\times10^{25} p_j$ | $\left(0.071\, f_j + 0.067 f_{j+1}\right) p_j$ |

Note 1: $f_j$ in vmr and $p_j$ in hPa.
Note 2: $f_j$ in ppmv and $p_j$ in hPa.

# Appendix B
# Content of the EOS MLS state vector

Section 3.5 discussed the construction and contents of the EOS MLS state vector. This appendix gives full details of all the components of the state vector currently envisaged. These details can be found in Table B.1. Some terms used in the table are defined here.

Each row of the table describes a separate component of the state vector. The first two columns describe the mathematical symbol and the the proposed name (for use in computer databases etc.) for each quantity. The meaning of the quantity is given in the third column. The fourth column describes the *type* of the quantity. The following types are used:

**Minor frame** quantities such as tangent pressures have one value or set of values for each EOS MLS minor frame.

**Orbital** quantities will typically be represented as some function (e.g. an interpolation) of a coordinate such as $\phi$, possibly with some additional long-term trend superimposed.

**Constant** quantities are literally constant in that a single value will be used for all profiles / major frames.

**Atmospheric profile** quantities describe vertical profiles of geophysical parameters on fixed pressure surfaces.

**Surface** quantities are simply atmospheric quantities on one particular pressure surface.

The fifth column describes the units for the quantity, and the last column describes the expected usage of the quantity.

**Retrieved** quantities are routinely retrieved in the MLS data processing (though not necessarily in all phases.)

**Constrained** quantities are, as the name implies, constrained to an *a priori* value. However, in most cases errors will not be propagated for these quantites.

**Characterized** quantities are constrained (as above) in routine processing. However, these quantities have sufficient impact on the direct measurements that their values could be retrieved in 'off-line' studies undertaken to characterize the instrument. These 'optimal' values can then be constrained in production processing.

Table B.1: Contents of EOS MLS state vector. See the text for an explanation of the terms and notes

| Symbol | Name | Meaning | Type | Units | Usage |
|---|---|---|---|---|---|
| *Tangent pressure terms* | | | | | |
| $\zeta^{118A}$ | PTAN-118A | R1A:118 radiometer tangent pressure. | Minor frame | $-\log_{10}$ [hPa] | Retrieved |
| $\zeta^{2T5H}$ | PTAN-2T5H | R5H:2T5 radiometer tangent pressure. | Minor frame | $-\log_{10}$ [hPa] | Retrieved |
| *Angular offset terms* | | | | | |
| $\delta^{ACS}$ | ACS-OFF | Angular offset between MLS and spacecraft attitude system. | Orbital | radians | Characterized |
| $\delta^{118A\rightarrow 118B}$ | ELEV-118B | Angular offset between R1A:118 and R1B:118 radiometers | Constant | radians | Characterized |
| $\delta^{118A\rightarrow 190}$ | ELEV-190 | Angular offset between R1A:118 and R2:190 radiometers | Constant | radians | Characterized |
| $\delta^{118A\rightarrow 240}$ | ELEV-240 | Angular offset between R1A:118 and R3:240 radiometers | Constant | radians | Characterized |
| $\delta^{118A\rightarrow 640}$ | ELEV-640 | Angular offset between R1A:118 and R4:640 radiometers | Constant | radians | Characterized |
| $\delta^{GHz\rightarrow THz}$ | ELEV-THZ | Angular offset between MLS GHz and THz modules | Orbital | radians | Probably retrieved |
| $\delta^{2T5H\rightarrow 2T5V}$ | ELEV-2T5V | Angular offset between R5H:2T5 and R5V:2T5 radiometers | Constant | radians | Characterized |
| *Geopotential height terms* | | | | | |
| $Z_0$ | Z0 | Geopotential height of reference pressure surface | surface | km | Retrieved |
| $Z_0^{118A}$ | Z0-118A | Geopotential height of reference pressure surface as determined by R1A:118 | surface | km | Retrieved |
| $Z_0^{118B}$ | Z0-118B | Geopotential height of reference pressure surface as determined by R1B:118 | surface | km | Retrieved |
| $Z_0^{240}$ | Z0-240 | Geopotential height of reference pressure surface as determined by R3:240 | surface | km | Retrieved |
| $Z_0^{2T5H}$ | Z0-2T5H | Geopotential height of reference pressure surface as determined by R5H:2T5 | surface | km | Retrieved |
| $Z_0^{2T5V}$ | Z0-2T5V | Geopotential height of reference pressure surface as determined by R5V:2T5 | surface | km | Retrieved |

*continued on next page . . .*

... continued from previous page

| Symbol | Name | Meaning | Type | Units | Usage |
|---|---|---|---|---|---|
| *Geometric terms* | | | | | |
| $\mathbf{v}^{\oplus}$ | VEL-EARTH | Component of air velocity at tangent point wrt. satellite due to Earth's rotation. | Atmospheric profile | m/s | Constrained |
| $\mathbf{v}^{S}$ | VEL-SAT | Component of air velocity at tangent point wrt. satellite due to satellite orbit. | Atmospheric profile | m/s | Constrained |
| $\mathbf{v}^{W}$ | VEL-WIND | Component of air velocity at tangent point wrt. satellite due to wind. | Atmospheric profile | m/s | Constrained |
| $R^{s}$ | SAT-GEOC-HT | Geocentric altitude of satellite | Geometric | km | Constrained |
| *Temperature terms* | | | | | |
| $\mathbf{T}$ | TEMP | Temperature | Atmospheric profile | K | Retrieved |
| $\mathbf{T}^{118A}$ | TEMP-118A | Temperature as determined by R1A:118 | Atmospheric profile | K | Retrieved |
| $\mathbf{T}^{118B}$ | TEMP-118B | Temperature as determined by R1B:118 | Atmospheric profile | K | Retrieved |
| $\mathbf{T}^{240}$ | TEMP-240 | Temperature as determined by R3:240 | Atmospheric profile | K | Retrieved |
| $\mathbf{T}^{2T5H}$ | TEMP-2T5H | Temperature as determined by R5H:2T5 | Atmospheric profile | K | Retrieved |
| $\mathbf{T}^{2T5V}$ | TEMP-2T5V | Temperature as determined by R5V:2T5 | Atmospheric profile | K | Retrieved |
| *Atmospheric composition terms* | | | | | |
| $\mathbf{f}_{H_2O}$ | H2O | Water vapor mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{H_2O}^{118}$ | H2O-118 | Water vapor mixing ratio as determined by R1A:118 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{H_2O}^{190}$ | H2O-190 | Water vapor mixing ratio as determined by R1A:190 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{H_2O}^{240}$ | H2O-240 | Water vapor mixing ratio as determined by R2:240 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{H_2O}^{640}$ | H2O-640 | Water vapor mixing ratio as determined by R4:640 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{H_2O}^{2T5H}$ | H2O-2T5H | Water vapor mixing ratio as determined by R5H:2T5 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{H_2O}^{2T5V}$ | H2O-2T5V | Water vapor mixing ratio as determined by R5H:2TV | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{HNO_3}$ | HNO3 | Nitric acid mixing ratio | Atmospheric profile | mixing ratio | Retrieved |

continued on next page ...

... continued from previous page

| Symbol | Name | Meaning | Type | Units | Usage |
|---|---|---|---|---|---|
| *Atmospheric composition terms (cont.)* | | | | | |
| $\mathbf{f}_{O_3}$ | O3 | Ozone mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{O_3}^{190}$ | O3-190 | Ozone mixing ratio as determined by R2:190 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{O_3}^{240-7}$ | O3-240-4 | Ozone mixing ratio as determined by R3:240.B7:O3 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{O_3}^{240-9}$ | O3-240-9 | Ozone mixing ratio as determined by R3:240.B9:O3 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{O_3}^{640}$ | O3-640 | Ozone mixing ratio as determined by R4:640 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{HCl}$ | HCl | Hydrogen chloride mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{ClO}$ | ClO | Chlorine monoxide mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{ClO}^{190}$ | ClO-190 | Chlorine monoxide mixing ratio as determined by R3:190 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{ClO}^{640}$ | ClO-640 | Chlorine monoxide mixing ratio as determined by R4:640 | Atmospheric Profile | mixing ratio | Retrieved |
| $\mathbf{f}_{N_2O}$ | N2O | Nitrous oxide mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{N_2O}^{190}$ | N2O-190 | Nitrous oxide mixing ratio as determined by R3:190 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{N_2O}^{640}$ | N2O-640 | Nitrous oxide mixing ratio as determined by R4:640 | Atmospheric Profile | mixing ratio | Retrieved |
| $\mathbf{f}_{OH}$ | OH | Hydroxyl mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{OH}^{2514H}$ | OH-2514H | Hydroxyl mixing ratio as determined by R5H:2T5.B15F:OH | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{OH}^{2514V}$ | OH-2514V | Hydroxyl mixing ratio as determined by R5V:2T5.B18F:OH | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{OH}^{2510H}$ | OH-2510H | Hydroxyl mixing ratio as determined by R5H:2T5.B16F:OH | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{OH}^{2510V}$ | OH-2510V | Hydroxyl mixing ratio as determined by R5V:2T5.B19F:OH | Atmospheric profile | mixing ratio | Retrieved |

continued on next page ...

... continued from previous page

| Symbol | Name | Meaning | Type | Units | Usage |
|---|---|---|---|---|---|
| *Atmospheric composition terms (cont.)* | | | | | |
| $\mathbf{f}_{HO_2}$ | HO2 | Hydroperoxy radical mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}^{650}_{HO_2}$ | HO2-650 | Hydroperoxy radical mixing ratio as determined by R4:640.B28M:HO2 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}^{660}_{HO_2}$ | HO2-660 | Hydroperoxy radical mixing ratio as determined by R4:640.B30M:HO2 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{Bro}$ | BrO | Bromine monoxide mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}^{625}_{BrO}$ | BrO-625 | Bromine monoxide mixing ratio as determined by R4:640.B30M:BrO | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}^{650}_{BrO}$ | BrO-650 | Bromine monoxide mixing ratio as determined by R4:640.B30M:BrO | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{CO}$ | CO | Carbon monoxide mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{HOCl}$ | HOCl | Hypochlorous acid mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{HCN}$ | HCN | Hydrogen cyanide mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}_{SO_2}$ | SO2 | Sulphur dioxide mixing ratio | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}^{190}_{SO_2}$ | SO2-190 | Sulphur dioxide mixing ratio as determined by R3:190 | Atmospheric profile | mixing ratio | Retrieved |
| $\mathbf{f}^{640}_{SO_2}$ | SO2-640 | Sulphur dioxide mixing ratio as determined by R4:640 | Atmospheric Profile | mixing ratio | Retrieved |
| *Mixing ratios for other trace species, minor isotopes of target gasses and excited states will also be included as necessary.* | | | | | |
| *Spectroscopic parameters.* | | | | | |
| $\mathbf{r}_i$ | ISOTOPE-R | Isotope ratios for some gasses | Profile / constant | dimensionless | Characterized |
| $\nu_0$ | LINE-FREQ | Spectral line center frequencies | Constant | MHz | Characterized |
| $\mathbf{w}_{cj}$ | P-BROAD | Pressure broadening parameters | Constant | MHz/hPa | Characterised |
| *Instrumental parameters* | | | | | |
| $\mathbf{r}_{sb}$ | SBR | Sideband ratios for each channel. | Constant | dimensionless | Characterized |
| $\mathbf{v}_c$ | CHAN-FREQ | Channel center frequencies for each channel. | Constant | MHz | Characterized |

# Appendix C
# Details of formulae used in this document.

## C.1 Calculus of vectors and matrices

Many of the manipulations in retrieval theory involve differentiating an expression with respect to a vector quantity. If the expression is a scalar, a vector of derivatives is obtained, each element representing the derivative of the scalar with respect to the corresponding element of the vector, as in

$$\left[\frac{\partial a}{\partial \mathbf{x}}\right]_i = \frac{\partial a}{\partial x_i}. \tag{C.1}$$

The derivative of a vector quantity with respect to another vector quantity can be represented by a matrix, according to

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = D_{ij} = \frac{\partial y_i}{\partial x_j}. \tag{C.2}$$

One can also construct vector equivalents for the various rules commonly associated with scalar calculus. In the following derivations consider vectors $\mathbf{x}$ and $\mathbf{y}$, with $\mathbf{y}$ depending on $\mathbf{x}$ according to $\mathbf{D} = \partial\mathbf{y}/\partial\mathbf{x}$. Many of the retrieval theory calculations involve evaluating expressions such as

$$\frac{\partial}{\partial \mathbf{x}}\mathbf{A}\mathbf{y}, \tag{C.3}$$

where $\mathbf{A}$ is a constant matrix. In order to evaluate these, it is necessary to consider individual components of the result.

$$\left[\frac{\partial}{\partial \mathbf{x}}\mathbf{A}\mathbf{y}\right]_{ij} = \frac{\partial}{\partial x_j}\sum_k A_{ik}y_k = \sum_k A_{ik}\frac{\partial y_k}{\partial x_j} = \sum_k A_{ik}D_{kj} = [\mathbf{AD}]_{ij}\,. \tag{C.4}$$

Another identity, involving the derivative of a scalar quantity with respect to a vector, is also common in retrieval problems. Again, the solution is easily found by considering components:

$$\left[\frac{\partial}{\partial \mathbf{x}}\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y}\right]_i = \frac{\partial}{\partial x_i}\sum_{jk} y_j A_{jk} y_k = \sum_j \frac{\partial y_j}{\partial x_i}A_{jk}y_k + y_j A_{jk}\frac{\partial y_k}{\partial x_i}. \tag{C.5}$$

A little thought is required before this expression can be recast into a matrix; several possible expressions result, one of which is

$$\frac{\partial}{\partial \mathbf{x}}\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{y} = \mathbf{D}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{y} + \mathbf{D}^{\mathrm{T}}\mathbf{A}\mathbf{y}. \tag{C.6}$$

## C.2 Details of the incremental information content calculation

Section 4.4.3 introduced the concept of 'information content' of a system, being related to the determinant of the covariance matrix. Computing the determinant of a matrix is an inherently unstable calculation[1]; instead a method is derived here which computes the information content by considering the incremental improvements made by introducing individual measurements one by one.

The incremental information content calculation is described in Rodgers 1996; the derivation is summarized here for clarity, and to set it in the context of the MLS case.

Equation 3.14 gave the covariance of the retrieved state vector as $\mathbf{S}_\mathbf{x} = \left[\mathbf{S}_\mathbf{a}^{-1} + \sum_i \mathbf{K}_i^\mathrm{T}\mathbf{S}_i^{-1}\mathbf{K}_i\right]^{-1}$. In this case we consider systems with only one measurement vector; this gives the covariance as

$$\mathbf{S}_\mathbf{x} = \left[\mathbf{S}_\mathbf{a}^{-1} + \mathbf{K}^\mathrm{T}\mathbf{S}_\mathbf{y}^{-1}\mathbf{K}\right]^{-1}. \tag{C.7}$$

It can be shown (through a somewhat complex series of manipulations, see Rodgers [in preparation]) that this is equivalent to

$$\mathbf{S}_\mathbf{x} = \mathbf{S}_\mathbf{a} - \mathbf{S}_\mathbf{a}\mathbf{K}^\mathrm{T}\left[\mathbf{K}\mathbf{S}_\mathbf{a}\mathbf{K}^\mathrm{T} + \mathbf{S}_\mathbf{y}\right]^{-1}\mathbf{K}\mathbf{S}_\mathbf{a}. \tag{C.8}$$

Now consider the case where the measurement covariance matrix $\mathbf{S}_\mathbf{y}$ is diagonal, and the measurements are entered sequentially as scalar values with variances $\sigma_j^2$. In this case, Equation C.8 can become an iterative expression

$$\begin{aligned}\mathbf{S}_\mathbf{x}^{(j)} &= \mathbf{S}_\mathbf{x}^{(j-1)} - \frac{\mathbf{S}_\mathbf{x}^{(j-1)}\mathbf{k}_j^\mathrm{T}\mathbf{k}_j\mathbf{S}_\mathbf{x}^{(j-1)}}{\mathbf{k}_j\mathbf{S}_\mathbf{x}^{(j-1)}\mathbf{k}_j^\mathrm{T} + \sigma_j^2} \\ &= \mathbf{S}_\mathbf{x}^{(j-1)}\left[\mathbf{I}_n - \frac{\mathbf{k}_j^\mathrm{T}\mathbf{k}_j\mathbf{S}_\mathbf{x}^{(j-1)}}{\mathbf{k}_j\mathbf{S}_\mathbf{x}^{(j-1)}\mathbf{k}_j^\mathrm{T} + \sigma_j^2}\right],\end{aligned} \tag{C.9}$$

where $\mathbf{S}_\mathbf{x}^j$ is the covariance of the solution after introducing measurement $j$ with $\mathbf{S}_\mathbf{x}^{(0)} = \mathbf{S}_\mathbf{a}$, and $\mathbf{k}_j$ is the weighting function for the $j$th measurement, i.e. the $j$th row of $\mathbf{K}$.

The additional information contributed by measurement $j$ is from Equation 4.18 given by

$$\delta H_j = \frac{1}{2}\log_2\left|\mathbf{S}_\mathbf{x}^{(j-1)}\right| - \frac{1}{2}\log_2\left|\mathbf{S}_\mathbf{x}^{(j)}\right|, \tag{C.10}$$

which, when applying the identities $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$ and $\left|\mathbf{I} + \mathbf{a}\mathbf{b}^\mathrm{T}\right| = 1 + \mathbf{b}^\mathrm{T}\mathbf{a}$, where $\mathbf{a}$ and $\mathbf{b}$ are column vectors, gives

$$\delta H_j = -\frac{1}{2}\log_2\left[1 - \frac{\mathbf{k}_j^\mathrm{T}\mathbf{S}_\mathbf{x}^{(j-1)}\mathbf{k}_j}{\mathbf{k}_j^\mathrm{T}\mathbf{S}_\mathbf{x}^{(j-1)}\mathbf{k}_j + \sigma_j^2}\right] = \frac{1}{2}\log_2\left[1 + \frac{\mathbf{k}_j^\mathrm{T}\mathbf{S}_\mathbf{x}^{(j-1)}\mathbf{k}_j}{\sigma_j^2}\right]. \tag{C.11}$$

This calculation is significantly more stable and also more efficient than computing the full determinant of the solution covariance.

[1] The determinant is the product of the eigenvalues of the matrix, consider the case of a 100×100 matrix whose eigen values are $\sim 10^{-6}$. This would have a determinant of $10^{-600}$ — too small to be represented on most computer architectures.

# Appendix D
# Notation conventions.

While the use of a consistant notation convention is desirable, it should not be achieved at the expense of reduced clarity. It would be inappropriate, for example, to use any character other than $g$ to describe the acceleration due to gravity. Hence, while there are exceptions to these conventions, in general the rules hold throughout the document.

**Scalars:** Scalars are represented by italic characters, e.g. $\alpha, i, n, M$.

**Vectors:** Vectors are shown as bold lower case characters, such as $\mathbf{x}$. To describe individual elements of a vector, the corresponding italic character is subscripted, so $[\mathbf{x}]_i = x_i$. Where the bold character is subscripted with an italic index, this indicates a specific vector from a set of vectors. So $\mathbf{y}_i$ is the $i$'th measurement vector.

**Matrices:** Bold upper case characters indicate matrices (e.g. $\mathbf{A}$). Again, subscripts on corresponding italic characters indicate individual elements, so $[\mathbf{A}]_{ij} = A_{ij}$. Also, as before, where the bold character is subscripted, this indicates a particular matrix in a family of matrices (so $\mathbf{K}_i$ is the $i$'th weighting function matrix.) In the case of covariance matrices, bold subscripts are used to indicate the covariance of a particular vector, thus $\mathbf{S}_\mathbf{x}$ is the covariance of $\mathbf{x}$ ($\mathbf{S}_i$ is a shorthand for $\mathbf{S}_{\mathbf{y}_i}$.)

**Subscripts:** In order to improve clarity, latin characters will typically be used to subscript quantities in state space, with Greek characters subscripting measurement space quantities, thus $K_{\alpha i} = \partial y_\alpha / \partial x_i$.

**Iterative processes:** In iterative process, the value of a quantity for a particular iteration is indicated by a parenthetical superscript, thus $\mathbf{x}^{(r+1)} = \mathbf{x}^{(r)} + \ldots$.

**Minor frame quantities:** It can often be useful to distinguish 'atmospheric' quantities from 'instrumental' ones. In most cases the atmospheric quantities will be represented by profiles on fixed pressure surfaces. Instrumental quantities are typically dependent on the minor frame (i.e. radiance integration period). Where such distinction is useful, minor frame quantities are indicated by an arrow placed over the symbol. For example the tangent point pressure for minor frame $i$ is indicated by $\vec{\zeta}_i$.

**Subvectors and submatrices:** Much of this work deals with subsections of matrices and vectors. These are indicated with bracketed subscripts. Examples of these are $\mathbf{x}_{[j]}$, the $j$'th subvector of $\mathbf{x}$, and $\mathbf{K}_{[\alpha i]}$ is the matrix $\partial \mathbf{y}_{[\alpha]} / \partial \mathbf{x}_{[i]}$

**Scaled quantities:** The $\tilde{}$ symbol is used to indicate quantities that have been scaled for numerical stability. Thus $\tilde{\mathbf{x}}$ represents the scaled state vector.

**Linerization points:** Where a linearization point has been chosen for the system the $\star$ superscript is used to indicate the use of linearized values. So, for example $\mathbf{y} \simeq \mathbf{y}^\star + \mathbf{K}^\star [\mathbf{x} - \mathbf{x}^\star]$.

# Appendix E
# EOS MLS signal designation nomenclature

## E.1 Motivation

The EOS MLS instrument contains seven radiometers observing five different spectral regions. The signals from these radiometers are subdivided into multiple bands, each of which is observed by a different spectrometer. The instrument contains a switch network which allows most spectral bands to be observed by one of two different spectrometers (or in many cases, both simultaneously). This switch network is present to provide both flexibility for power saving modes of instrument operation, and some redundancy.

The complexity of this system is such that a complete nomenclature scheme is essential for a clear understanding of the instrument. Such a system has been devised with the intention that it will be used in all the aspects of the instrument and software, from hardware drawings to science data processing software.

## E.2 The nomenclature scheme

The scheme consists of up to five fields, in the form

<Radiometer>.<Band>.<Switch>.<Spectrometer>.<Channel>

Such a specification has many useful properties; in particular, fields can be ignored if they are not relevant to the specification. For example, in the Level 2 software, the user could specify that ozone is to be retrieved from `R2:190.B6F:O3`, without needing to specify a switch and/or spectrometer, as either of the two alternatives will be appropriate.

Earlier fields can also be ignored. For example, the instrument command and data handling system will typically only consider the channel and spectrometer fields, as the switch, band and radiometer information are of little relevance to instrument data handling activities.

The following subsections explain each of the fields in the specification.

### E.2.1 Radiometers

The instrument consists of seven radiometers, measuring five different spectral regions. These are called `R1A`, `R1B`, `R2`, `R3`, `R4`, `R5H` and `R5V`. `R1A` and `R1B` are redundant 118 GHz radiometers. `R5H` and `R5V` are alternate polarizations of the 2.5 THz signal. As a courtesy to the reader, the radiometer field can also contain an indication of the frequency, separated from the number by a colon. Thus the full names for the radiometers are: `R1A:118`, `R1B:118`, `R2:190`, `R3:240`, `R4:640`, `R5H:2T5`, and `R5V:2T5`. The frequency information can be omitted, but if present it must be correct.

### E.2.2 Bands

The MLS spectral bands are numbered sequentially throughout the instrument; the numbering does not restart at one for each new radiometer. The band specification begins with the letter `B`, followed by the number of the band. Following that, there is an optional character indicating whether the primary signal is in the upper (`U`) or lower (`L`) sideband of the radiometer. Following this optional character is a compulsory one indicating the type of spectrometer used for this band. Thus, `F` indicates a standard 25 channel filter bank, `M` a 'mid-band' 11 channel filter bank, `D` a digital autocorrelation spectrometer, and `W` a set of four individual wide filters.

Following this (separated by a colon) there can be additional courtesy information to the user describing the primary target of the band (e.g. `O3` for ozone, `PT` for temperature/pressure.)

### E.2.3 Switch

Nearly all of the MLS spectral bands can be routed to one or two different spectrometers through a switch network. Most bands have a direct route to a spectrometer through no switch; this is designated by `S0`. Alternatively, most bands can be routed to an alternative spectrometer through one of five switches; such cases are designated `S1`$\cdots$`S5`.

### E.2.4 Spectrometer

Following the switch field, the spectrometer type and number is indicated. The instrument contains four types of spectrometers. There are nineteen 'standard' 25 channel filter banks, designated `FB25-1` $\cdots$ `FB25-19`; five 'mid-band' 11 channel filter banks (`MB11-1` $\cdots$ `MB11-5`); four digital autocorrelation spectrometers (`DACS-1` $\cdots$ `DACS-4`); and three sets of 4 individual wide filters (`WF4-1` $\cdots$ `WF4-3`.)

### E.2.5 Channels

Channels are simply specified by a `C`, followed by a number. Channels are numbered from one, except in the DACS where the numbering starts from zero.

### E.2.6 General comments

This nomenclature system deliberately contains much redundancy designed to improve clarity; however, it may be trimmed down if the user wishes.

- Radiometers may be specified without their frequency information.
- Bands may be specified without their intended target, or upper/lower sideband information.
- The switch and spectrometer information are redundant; thus one may be omitted if desired.
- If a band is specified the radiometer specification is redundant and may be omitted. This is discouraged, however, as clarity is lost.

## E.3 The valid MLS signals

Tables E.1 and E.2 list all the various radiance signals that can be measured by EOS MLS.

**Table E.1:** The nominal MLS measurement set.

| Radiometer | Band | Switch | Spectrometer | Channels |
|---|---|---|---|---|
| R1A:118 | B1F:PT | S0 | FB25-1 | C1···C25 |
| R2:190 | B2F:H2O | S0 | FB25-2 | C1···C25 |
| R2:190 | B3F:N2O | S2 | FB25-3 | C1···C25 |
| R2:190 | B4F:HNO3 | S0 | FB25-4 | C1···C25 |
| R2:190 | B5F:CLO | S0 | FB25-5 | C1···C25 |
| R2:190 | B6F:O3 | S0 | FB25-6 | C1···C25 |
| R3:240 | B7F:O3 | S0 | FB25-7 | C1···C25 |
| R3:240 | B8F:PT | S3 | FB25-8 | C1···C25 |
| R3:240 | B9F:CO | S0 | FB25-9 | C1···C25 |
| R4:640 | B10F:CLO | S0 | FB25-10 | C1···C25 |
| R4:640 | B11F:BRO | S0 | FB25-11 | C1···C25 |
| R4:640 | B12F:N2O | S4 | FB25-12 | C1···C25 |
| R4:640 | B13F:HCL | S0 | FB25-13 | C1···C25 |
| R4:640 | B14F:O3 | S0 | FB25-14 | C1···C25 |
| R5H:2T5 | B15F:OH | S5 | FB25-15 | C1···C25 |
| R5H:2T5 | B16F:OH | S0 | FB25-16 | C1···C25 |
| R5H:2T5 | B17F:PT | S0 | FB25-17 | C1···C25 |
| R5V:2T5 | B18F:OH | S0 | FB25-18 | C1···C25 |
| R5V:2T5 | B19F:OH | S0 | FB25-19 | C1···C25 |
| R1A:118 | B22D:PT | S1 | DACS-1 | C0···C128 |
| R2:190 | B23D:H2O | S0 | DACS-2 | C0···C128 |
| R3:240 | B24D:O3 | S0 | DACS-3 | C0···C128 |
| R3:240 | B25D:O3 | S0 | DACS-4 | C0···C128 |
| R2:190 | B27M:HCN | S0 | MB11-1 | C1···C11 |
| R4:640 | B28M:HO2 | S0 | MB11-2 | C1···C11 |
| R4:640 | B29M:HOCL | S0 | MB11-3 | C1···C11 |
| R4:640 | B30M:HO2 | S0 | MB11-4 | C1···C11 |
| R4:640 | B31M:BRO | S0 | MB11-5 | C1···C11 |
| R1A:118 | B32W:PT | S0 | WF4-1 | C1···C4 |
| R3:240 | B33W:O3 | S0 | WF4-2 | C1···C4 |

**Table E.2:** The alternate MLS measurement set.

| Radiometer | Band | Switch | Spectrometer | Channels |
|---|---|---|---|---|
| R1A:118 | B1F:PT | S3 | FB25-8 | C1···C25 |
| R1B:118 | B21F:PT | S4 | FB25-12 | C1···C25 |
| R1B:118 | B21F:PT | S3 | FB25-8 | C1···C25 |
| R2:190 | B2F:H2O | S2 | FB25-3 | C1···C25 |
| R2:190 | B4F:HNO3 | S2 | FB25-3 | C1···C25 |
| R2:190 | B5F:CLO | S2 | FB25-3 | C1···C25 |
| R3:240 | B7F:O3 | S3 | FB25-8 | C1···C25 |
| R3:240 | B8F:PT | S2 | FB25-8 | C1···C25 |
| R3:240 | B9F:CO | S3 | FB25-8 | C1···C25 |
| R4:640 | B13F:HCL | S4 | FB25-12 | C1···C25 |
| R4:640 | B14F:O3 | S4 | FB25-12 | C1···C25 |
| R5H:2T5 | B16F:OH | S5 | FB25-15 | C1···C25 |
| R5H:2T5 | B17F:PT | S5 | FB25-15 | C1···C25 |
| R5V:2T5 | B18F:OH | S5 | FB25-15 | C1···C25 |
| R5V:2T5 | B19F:OH | S5 | FB25-15 | C1···C25 |
| R5V:2T5 | B20F:PT | S3 | FB25-8 | C1···C25 |
| R5V:2T5 | B20F:PT | S5 | FB25-15 | C1···C25 |
| R1B:118 | B26D:PT | S1 | DACS-1 | C0···C128 |
| R1B:118 | B34W:PT | S0 | WF4-3 | C1···C4 |

# Bibliography


Barath, F., M. Chavez, R. Cofield, D. Flower, M. Frerking, M. Gram, W. Harris, J. Holden, R. Jarnot, W. Kloezeman, G. Klose, G. Lau, M. Loo, B. Maddison, R. Mattauch, R. McKinney, G. Peckham, H. Pickett, G. Siebes, F. Soltis, R. Suttie, J. Tarsala, J. Waters, and W. Wilson (1993). The Upper Atmosphere Research Satellite Microwave Limb Sounder Experiment. *Journal of Geophysical Research 98*(D6), 10,751–10,762.

Data Assimilation Office (1996). Algorithm Theoretical Basis Document Version 1.01. Technical report, Data Assimilation Office, NASA Goddard Space Flight Center.

Dudhia, A. and N. Livesey (1995). Validation of the Improved Stratospheric and Mesospheric Sounder temperature measurements. *Journal of Geophysical Research 101*(D6), 9,795–9,809.

Filipiak, M. (1999). EOS MLS retrieved geophysical parameter precision estimates. version 1.1. Technical report, University of Edinburgh, Department of Meteorology.

Fishbein, E., R. Cofield, L. Froidevaux, R. Jarnot, T. Lungu, W. Read, Z. Shippony, J. Waters, I. McDermid, T. McGee, U. Singh, M. Gross, A. Hauchecorne, P. Keckhut, M. Gelman, and R. Nagatani (1996). Validation of UARS Microwave Limb Sounder temperature and pressure measurements. *Journal of Geophysical Research 101*(D6), 9,983–10,016.

Fleagle, R. and J. Businger (1963). *An Introduction to Atmospheric Physics*. Academic Press.

Froidevaux, L., W. Read, T. Lungu, R. Cofield, E. Fishbein, D. Flower, R. Jarnot, B. Ridenoure, Z. Shippony, J. Waters, J. Margitan, I. McDermid, R. Stachnik, G. Peckham, G. Braathen, T. Deshler, J. F. D. Hoffmann, and S. Oltmans (1996). Validation of UARS Microwave Limb Sounder ozone measurements. *Journal of Geophysical Research 101*(D6), 10,017–10,060.

Golub, G. and C. VanLoan (1996). *Matrix Computations, Third edition*. Johns Hopkins University Press.

Hess, S. (Ed.) (1959). *An Introduction to Theoretical Meteorology*. Krieger Publishing Company. (Holt and Rinehart and Winston).

Jarnot, R. (1999). EOS MLS level 1 algorithm theoretical basis. version 1.1. Technical report, Jet Propulsion Laboratory. D-15210.

Livesey, N. et al. (1999). The UARS Microwave Limb Sounder version 5 dataset. In preparation.

Marks, C. and C. Rodgers (1993). A retrieval method for atmospheric composition from limb emission measurements. *Journal of Geophysical Research 98*(D8), 14,939–14,953.

Press, W., B. Flannery, S. Teukolsky, and W. Vetterling (1986). *Numerical Recipes — The Art of Scientific Computing*. Cambridge University Press.

Read, W., Z. Shippony, et al. (1999). EOS MLS forward model algorithm theoretical basis document. Technical report, Jet Propulsion Laboratory. In preparation.

Read, W., J. Waters, C. Smallcomb, E. Stone, Z. Shippony, D. Wu, A. Smedley, S. Oltmans, D. Kley, H. Smit, J. Mergenthaler, and M. Karki (1999). Mls uth validation. Manuscript in preparation.

Reber, C. (1993). The Upper Atmosphere Research Satellite (UARS). *Geophysical Research Letters 20*(12), 1,215–1,218.

Reber, C., C. Trevathan, R. McNeal, and M. Luther (1993). The Upper Atmosphere Research Satellite (UARS) mission. *Journal of Geophysical Research 98*(D6), 10,643–10,647.

Rodgers, C. (1976). Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Reviews of Geophysics and Space Physics 14*(4), 609–624.

Rodgers, C. (1990). Characterisation and error analysis of profiles retrieved from remote sounding measurements. *Journal of Geophysical Research 95*(D5), 5,587–5,595.

Rodgers, C. (1996). Information content and optimisation of high spectral resolution measurements. In P. Hays and J. Wang (Eds.), *Optical Spectroscopic Techniques and Instrumentation for Atmospheric and Space Research II*, pp. 136–147. SPIE.

Rodgers, C. (In preparation). Inverse methods for atmospheric sounding: Theory and practice. See `http://www.atm.ox.ac.uk/user/rodgers/`.

Taylor, F., C. Rodgers, J. Whitney, S. Werrett, J. Barnett, G. Peskett, P. Venters, J. Ballard, C. Palmer, R. Knight, P. Morris, and T. Nightingale (1993). Remote sensing of atmospheric structure and composition by pressure modulation radiometry from space: The ISAMS experiment on UARS. *Journal of Geophysical Research 98*(D6), 10,799–10,814.

Waters, J. (1999). An overview of the EOS MLS experiment. version 1.1. Technical report, Jet Propulsion Laboratory. D-15745.

Waters, J., W. Read, L. Froidevaux, T. Lungu, V. Perun, R. Stachnik, R. Jarnot, R. Cofield, E. Fishbein, D. Flower, J. Burke, J. Hardy, L. Nakamura, B. Ridenoure, Z. Shippony, and R. Thurstans (1996). Validation of UARS Microwave Limb Sounder ClO measurements. *Journal of Geophysical Research 101*(D6), 10,091–10,127.

This document was produced with LaTeX 2ε, using Times Roman and the *MathTime complete* font set. The figures were generated by xfig, the pstricks package and IDL. The pdf version of this document was generated by Adobe Acrobat. The hypertext links were automatically generated by the hyperref package.
The radiance data used in Figure 3.1 were provided by Mark Filipiak of the Meteorology Department, Edinburgh University, Scotland.